Synthesis and Investigation of Benzimidazole and Carbazole β-haematin Inhibiting Scaffolds with Antimalarial Activity

Fabrizio P. L'abbate

Thesis presented for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Chemistry

UNIVERSITY OF CAPE TOWN

February 2018

Supervisors: Professor Timothy J. Egan and Professor Roger Hunter

Declaration

I, Fabrizio P. L'abbate, hereby declare that all of the work presented in this thesis, "**Synthesis and Investigation of Benzimidazole and Carbazole β-haematin Inhibiting Scaffolds with Antimalarial Activity**", is my own data except for that which is properly acknowledged. I declare that no part of the above-titled thesis, is being, or has been submitted for another degree at this or any other university. I hereby authorise the University of Cape Town free license to reproduce this work, in whole or in part, for the purpose of research.

I hereby present this thesis for examination for the degree of Doctor of Philosophy (PhD) in Chemistry.

Signature: Signed by candidate **Date: 09/02/2018**

Abstract

Chloroquine was one of the main malarial treatments until the late 1960s when resistance began to emerge. This antimalarial targets haemozoin formation which causes a cytotoxic accumulation of free haem in the malaria parasite leading to parasite death. This is still one of the most promising pathways for treatment of the most prevalent species of malaria parasite, *Plasmodium falciparum* to date but, owing to growing resistance to chloroquine and other current antimalarial drugs, there is a dire need for new drugs. One strategy is to investigate non-chloroquine haemozoin inhibitors. High-throughput screening (HTS) was previously used to investigate novel β-haematin (synthetic haemozoin) inhibitors with promising *P. falciparum* growth inhibition activities. Of the 144 330 compounds screened, two hit compounds were selected for investigation in this project with two different scaffolds, namely benzimidazole and carbazole indole.

In order to preselect benzimidazole derivatives for synthesis, Discovery Studio and Pipeline Pilot where used in tandem to enumerate 325 728 *in silico* compounds. These were filtered according to predicted β-haematin inhibition activities, followed by predicted malaria parasite growth activities using previously developed models based on Bayesian statistics. The predicted active compounds were further subjected to an *in silico* aqueous solubility model and separated according to predicted solubility values however, only 68 out of the 35 124 active compounds showed moderate solubility whilst the rest were poorly soluble. From this data, eighteen compounds were chosen for synthesis with varying functional groups. Using the same Bayesian models, biological activities for seven fragment compounds derived from the benzimidazole hit compound were predicted. Six out of seven were predicted to be β-haematin inhibitors while five out of seven were predicted active against the malaria parasite growth inhibition model. Similar Bayesian predictions were carried out on the seven proposed carbazole indole compounds with three compounds predicted to be β-haematin inhibitors while six compounds were predicted to be active against the malaria parasite growth inhibition model.

The eighteen benzimidazole compounds were synthesized using a two-step synthesis, via a condensation reaction using polyphosphoric acid (PPA), 4-aminobenzoic acid and *o*-phenylenediamine to form the primary amine benzimidazole intermediate after which an

acylation reaction with the appropriate acid chloride furnished the desired compounds. β-haematin inhibition analysis revealed a 78% hit rate compared to the Bayesian predictions which resulted in a 24-fold enrichment compared to random screening. SAR analysis revealed an activity trend related to the position of substituents on the ring system as follows: *para* < *ortho* < *meta*. The type of ring system was also investigated, with a trend of phenyl < furan < pyrrole < thiophene < pyridyl found. The fragment compounds were either purchased or synthesized via standard acylation conditions using acid chlorides or acetic anhydride with primary amines as before. β-haematin inhibition analysis showed all these compounds to be inactive at the 100 μM cut-off but these compounds were still carried through to the next stage of testing in spite of these results.

Molecular docking was carried out on all eighteen benzimidazole compounds in Materials Studio using the (001) and (011) β-haematin crystal faces for adsorption, together with a modified CVFF force-field. This showed a correlation between adsorption energies of the (011) β-haematin crystal face with the experimental β-haematin inhibition values. This indicated that the (011) β-haematin crystal face was the most important for β-haematin inhibition. Analysis of the benzimidazole compounds and their π-π and hydrogen bonding interactions was performed. The number of π-π interactions were found to be important for β-haematin inhibition activity. Both sets of benzimidazole compounds were tested against the NF54 chloroquine sensitive malaria parasite using growth inhibition assays with a 50% hit rate shown for the benzimidazole compounds and a 71% hit rate for the fragment study leading to a 26-fold and 36-fold enrichments compared to random screening. SAR analysis of the benzimidazole compounds revealed a trend for activity in relation to substituent position of *para* ≈ *ortho* < *meta* and a ring system trend of phenyl < pyridyl < thiophene < furan < pyrrole. The benzimidazole compounds were further tested against the chloroquine resistant Dd2 *P*. *falciparum* strain which showed that disubstituted compounds were more active against this strain. Cellular haem fractionation studies revealed an increase in free haem and decrease in haemozoin confirming that haemozoin inhibition is the mode of action for the benzimidazole compounds. QSAR analysis of these compounds revealed a correlation between the -Log(*P*. *falciparum* IC_{50}) which is also known as pLog(*P*. *falciparum* IC_{50}) and 1/β-haematin IC_{50}, number of hydrogen bond donors and molecular depth with 1/β-haematin IC_{50} the most dominant term.

The first four carbazole indole compounds were synthesized using a two-step synthesis via deprotonation of carbazole and reaction with epichlorohydrin or 1,3-dibromopropane to furnish the epoxide or alkylbromine intermediates. These intermediates underwent a further S_N2 reaction using deprotonated indole to furnish four final compounds. Synthesis of another three derivatives required benzyl protection of 7-hydroxyindole alcohol first, followed by reaction with the epoxide intermediates via an S_N2 mechanism to furnish the final three compounds. Analysis using the turbidimetric solubility assay revealed the best aqueous solubility range of this series of compounds to be 10-20 µM (moderately soluble). β-haematin inhibition studies were carried out on this series of compounds with a 100% hit rate found when compared to the Bayesian model data which lead to 30-fold enrichment when compared to random screening. SAR analysis showed an increase in the number of hydroxyl groups led to an increase in β-haematin inhibition activity. Docking studies were performed on these seven compounds and showed that hydrogen bonding played a role in anchoring the molecules in the binding pocket on the crystal surface with increased adsorption energies seen with an increase in the number of hydroxyl groups. Malaria parasite growth inhibition studies showed no compounds to be active against the NF54 and Dd2 strains at the 2 µM cut-off. Cellular haem fractionation studies on the carbazole indole compounds showed that this series of compounds acts via a mechanism that results in inhibition of haemoglobin uptake into the food vacuole and not via haemozoin inhibition.

Acknowledgments

Many thanks to all the people who contributed to this work. Special thanks to the following people for numerous contributions to this project:

My supervisor, Prof. Timothy J. Egan, thank you for all the opportunities you have given me throughout the project, your vast knowledge pool that you have shared with me, your time, patience and encouragement.

My co-supervisor, Prof. Roger Hunter, thank you for challenging me to be a better organic chemist and allowing me to use all your facilities.

Dr. Wade Peterson, thank you for invaluable mentoring during the early years. Without it I would not be the organic chemist I am today.

Roxanne Mohunlal, for your support throughout all of my postgraduate studies, a great friendship and all the amazing shared experiences.

Dr. Shankari Nair, for being the best synthetic lab friend that I could always count on. You were always the voice of reason in difficult times.

Shakeela Sayed, I could always rely on you for help and support. Thank you for all the amazing talks and for all the motivation during the last stretch of this project.

To all the past and present members of Bioinorganic research group (Haem team), in particular Dr. John Woodland, Dr. Aneesa Omar, Nikki Kuter, Dr. David Kuter and Dr. John Okombo.

To all the past and present members of the Synthetic chemistry group, in particular Dr. Rudy Cozett, Dr. Ana Andrijevic, James Biwi and Daniel Kusza.

Dr. Jill Combrinck, for all your help with *in vitro* testing. You were an enormous help in this endeavour and your assistance will not be forgotten.

Dr. Katherine de Villiers and her research group at Stellenbosch University for their collaboration on the NIH project.

Prof. Peter Smith, at the Division of Pharmacology, UCT Medical School, for all allowing my compounds to be tested at your facilities.

Peter Roberts, for all the NMR collections.

The Mass Spectrometry Unit at Stellenbosch University and the Centre for High Performance Computing (CHPC) for their services.

The National Institutes of Health (NIH) for funding under grant no. R01AI110329. Further thanks to the University of Cape Town for funding.

To the love of my life, Lauren Kim Mc George, for all your love and support throughout this endeavour. You are my pillar of strength and without you, this project would not be possible.

To my parents, Giorgio and Renee L'abbate, for always wanting the best for me and supporting me even through the hard times. Your love and support made this project possible teaching me to persevere even through insurmountable odds.

This project is dedicated to my late, best friend, Derrick Kropman. Thank you for teaching me how to enjoy life and may you rest in peace.

Abbreviations

ACT	Artemisinin-based combination therapy
ADME	Absorption, distribution, metabolism and excretion
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate
d	Doublet
DCC	*N,N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
dd	Doublet of doublets
ddd	Doublet of doublets of doublets
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
dt	Doublet of triplets
ee	Enantiomeric excess
Eq.	Equivalents
ERG	Electron releasing group
EtOAc	Ethyl acetate
EWR	Electron withdrawing group
Fe(II)PPIX	Iron(II) protoporphyrin IX
Fe(III)PPIX	Iron(III) protoporphyrin IX
GSK	GlaxoSmithKline
HEPES	4-(2-Hydroxyethyl-1-Piperazineethanesulfonic acid
HOBt	1-Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
HTS	High throughput screening
IC_{50}	50% Inhibitory concentration
IR	Infrared
m	Multiplet
MeOH	Methanol
NHS	*N*-hydroxysuccinimide
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NP-40	Nonidet P-40
PBS	Phosphate-buffered saline
*Pf*CRT	*Plasmodium falciparum* chloroquine resistance transporter
*Pf*mdr1	*Plasmodium falciparum* multidrug resistance gene
PPA	Polyphosphoric acid
PXRD	Powder X-ray diffraction
RI	Resistance index
Rf	Retention factor
s	Singlet
SAR	Structure activity relationships
SDS	Sodium dodecyl sulphate
SI	Selectivity index
t	Triplet
THF	Tetrahydofuran
TLC	Thin layer chromatography
Triton X-100	Polyethylene glycol *tert*-octylphenyl ether
TWEEN® 20	Polyoxyethylenesorbitan monolaurate
TWEEN® 80	Polyethylene glycol sorbitan monolaurate
WHO	World Health Organisation

Table of Contents

Chapter 1: Introduction

1.1 Historical Outline of Malaria

Malaria is one of the oldest diseases known to man with records dating back from the sixth century BC described in Hindu texts. However, it was only in 400 BC that Hippocrates noticed enlarged spleens, malarial fevers and general poor health of people living in marsh lands. Due to this condition said to be originating from the marsh lands and swamps, over the next 2 500 years the name, spoiled air or mal'aria was developed by the Romans.[1,2] It was only in 1880 when Charles Louis Alphonse Laveran first described spherical parasites containing black pigment in red blood cells that significant interest in the malaria parasite began to develop.[1–4] Over more than a century there have been approximately 200 species of parasite that belong to the genus *Plasmodium* reported, with five of those infecting humans.[5,6] The most widespread, infective and fatal of these is *Plasmodium falciparum*.[6]

1.2 Current Malaria Scourge

The World Health Organisation (WHO)[7] malaria report describes the current malaria status of the world, using data collected from 91 malaria stricken countries. There were an estimated 216 million cases in 2016 with 80% of them occurring in sub-Saharan Africa. Artemisinin-based combination therapy (ACT) is the treatment of choice and has managed to decrease mortality rates with no resistance to the drug combination shown in Africa. Even though this treatment has substantially decreased the mortality rate, it was estimated that globally in 2016 malaria caused 445 000 deaths, with 91% of those deaths originating in sub-Saharan Africa. The main challenges facing malaria treatment are funding and emerging resistance to current therapies which encourages research into novel treatments. Figure 1.1 shows the distribution of malaria cases in the world comparing the year 2000 to 2016.

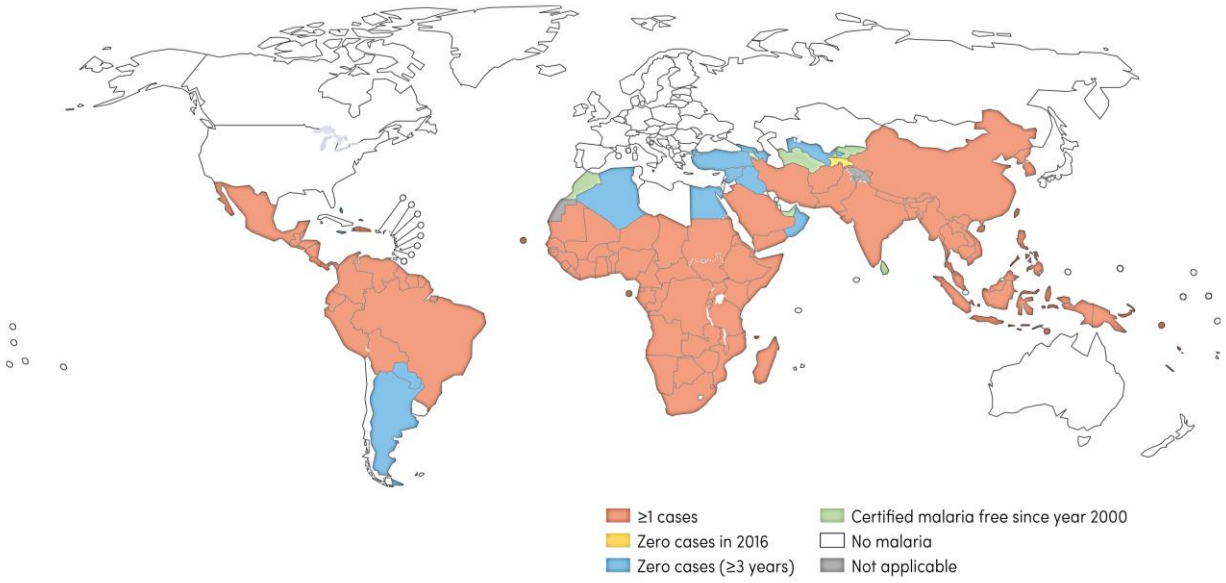


Figure 1.1: A world map showing countries and territories with cases of malaria in 2000 compared to 2016. The WHO can issue a certificate for a zone if it is malaria free for three consecutive years. These countries are coloured in blue while those that have seen zero cases in 2016 are coloured in yellow. Countries with one or more cases are coloured in red. Copyright © WHO malaria report 2017.[7]

1.3 The Malaria Parasite Life Cycle

It took almost 70 years to fully describe the malaria parasite life cycle starting from the first description in 1880 by Laveran.[2,3] The life cycle is similar in all five species of *Plasmodium* that infect humans. It is comprised of three stages, namely the liver, blood and mosquito stages which are shown in the life cycle of *P. falciparum* in Figure 1.2. It begins with the bite of a female *Anopheles* mosquito which transfers, with its saliva, motile sporozoites which enter the blood stream through the human dermis and quickly enter the liver through Kupffer cells, to begin the liver stage. In the liver they invade the liver hepatocytes and undergo asexual multiplication to form about 30 000 merozoites each. The blood stage begins after about a week, when the merozoites are released into the blood stream and infect the erythrocytes or red blood cells. During a 48 hour period, trophozoites develop from rings to schizonts which upon maturation, rupture releasing merozoites again. The development into schizonts is fuelled by haemoglobin degradation and at this stage the characteristic black pigment is seen in people without preimmunity. This process is repeated until the victim is treated or dies.

This is also the stage which gives rise to all the clinical symptoms seen in malaria patients. As infection increases some merozoites develop into male and female gametocytes in the red blood cells which circulate in the blood stream and are taken up by the female *Anopheles* mosquito. Inside the mosquito these gametocytes mature into gametes after which fertilization occurs and an ookinete is formed. This ookinete penetrates the gut wall and becomes an oocyst which multiplies and forms sporozoites which travel to the mosquito's salivary glands ready for infection into a human host.[2,6,8–10]

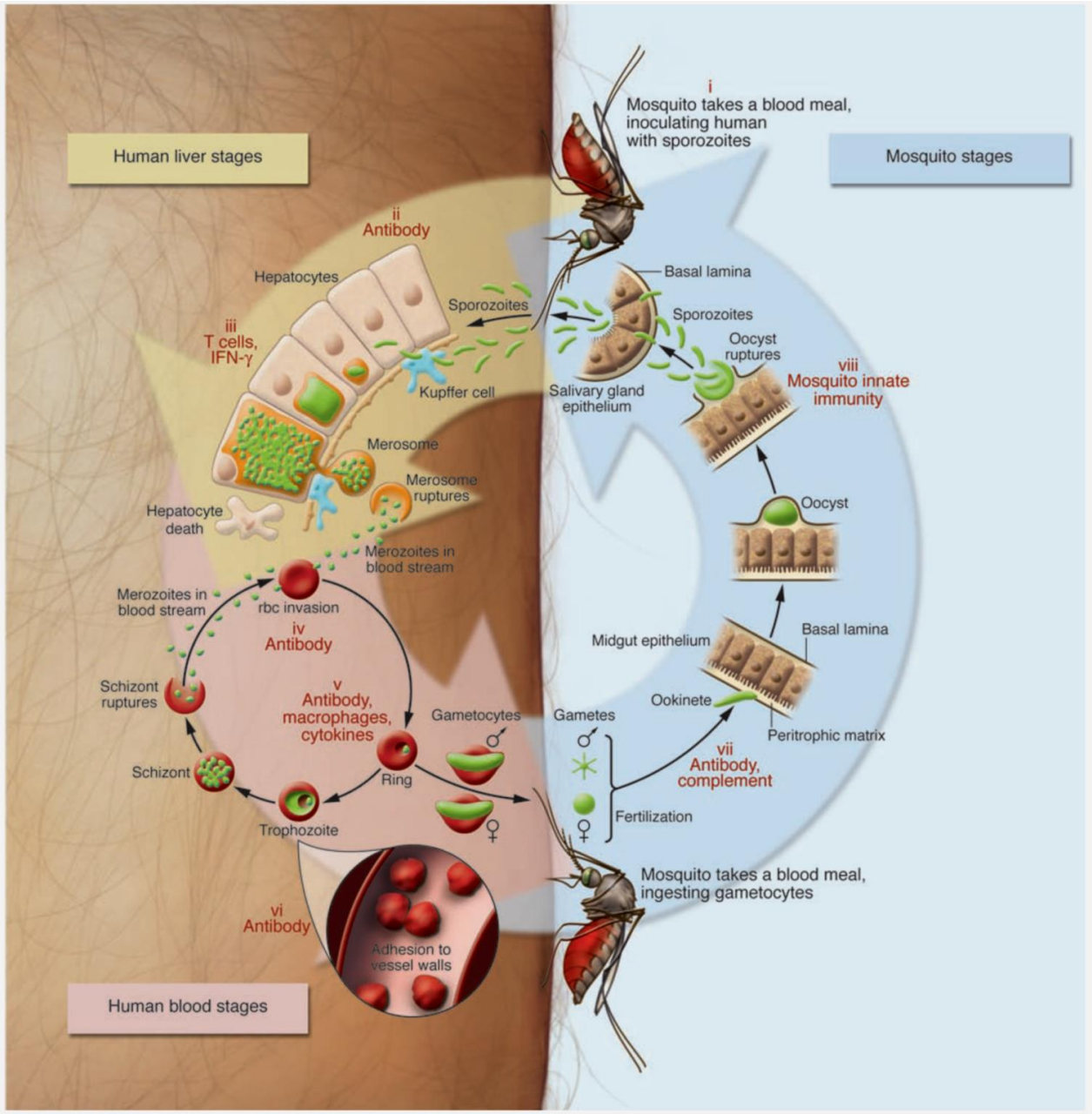


Figure 1.2: The life cycle of *Plasmodium falciparum*. Reprinted with permission from Greenwood *et al*.[8] Copyright © 2008 American Society for Clinical Investigation.

1.4 The Mechanism of Haemoglobin Degradation in *P. falciparum*

The malaria parasite ingests haemoglobin from the host cytoplasm by a process called endocytosis using a cytostome via which it is trafficked between the parasite and the red blood cell. The haemoglobin is transported to the acidic food vacoule (secondary lysosome), also known as the digestive vacuole, with a pH between 4.8 and 5.2 via vesicles.[6,11] The haemoglobin is digested by several proteases, plasmepsins I, II and histoaspartic protease releasing haem and globin. The globin is further degraded by plasmepsin IV, falcipains 2 and 3 and falcilysin which release peptides that are further broken down by aminopeptidases to amino acids, of which a fraction are used by the parasite.[12] These amino acids are used for maturation and growth of the parasite since the parasite has limited capabalities of synthesizing its own amino acids or acquiring them from elsewhere.[11] Some studies estimate that up to 80% of the host haemoglobin is degraded, but this is not solely to obtain amino acids as only about 15% of the amino acids are used while the rest are discharged to the surrounding plasma.[11–14] The other functions of haemoglobin degradation have been suggested to be to provide space for parasite growth and provide osmotic balance within the cell.[12] Another by-product of haemoglobin degradation is haem or ferroprotoporphyrin IX (Fe(II)PPIX) which is rapidly oxidised to ferriprotoporphyrin IX (Fe(III)PPIX) by oxygen in the acidic food vacoule. This is cytotoxic to the malaria parasite due to its capability to bring about lipid peroxidation especially in large amounts. Therefore the parasite has developed an ingenious method of detoxification by converting Fe(III)PPIX into an insoluble crystalline material called haemozoin via a biomineralisation process.[12,15] The complete haemoglobin degradation pathway is summarised in Figure 1.3.

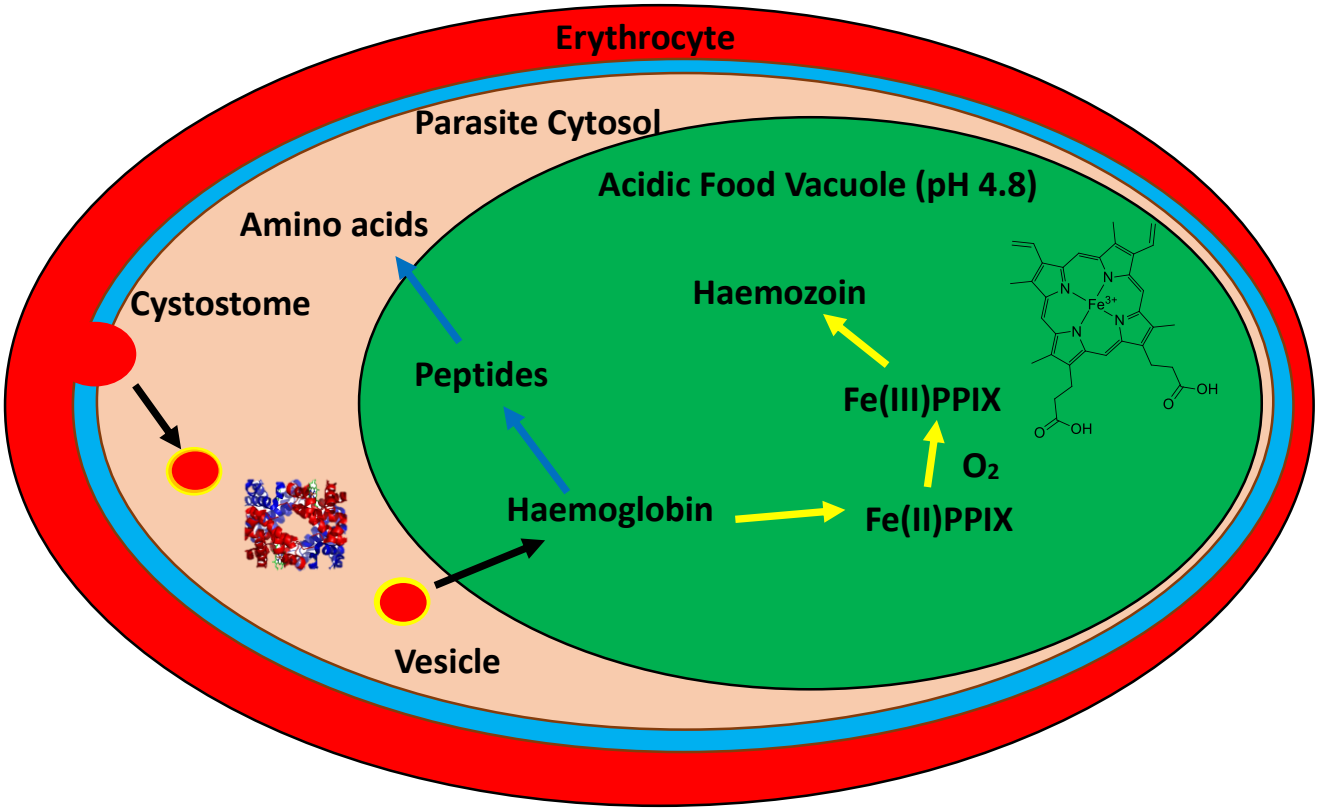


Figure 1.3: The degradation of haemoglobin in a parasitized red blood cell or erythrocyte resulting in the formation of amino acids and haemozoin.

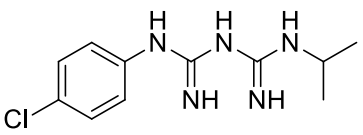
1.5 Chemotherapy of Malaria

Malaria chemotherapy currently relies on four major drug classes, namely, artemisinin and derivatives thereof, the antifolates, the antimicrobials and the quinoline derivatives.[9,16] These drug classes are discussed in detail in the next section.

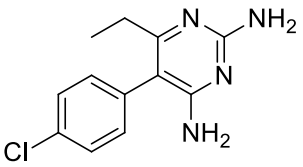
1.5.1 Antifolates and Antimicrobials

During the second world war, supplies of known antimalarial drugs at the time were low and thus there was a large research initiative started in the United Kingdom for discovery of a new drug. This led to the discovery of proguanil (chlorguanide) (Scheme 1.1) which is formed via opening the pyrimidine ring to form a biguanide. This compound was found not to be ideal for acute cases of malaria due to the slow action of the drug which was also found to induce resistance in the parasite. This success, however did start a research initiative into pyrimidine derivatives. In 1951, Falco and Hitchings developed pyrimethamine (Scheme 1.1), belonging to the 2,4-aminopyrimides class, which was hailed an important advancement due to its increased activity compared to proguanil.[17] However, as with proguanil, resistance to

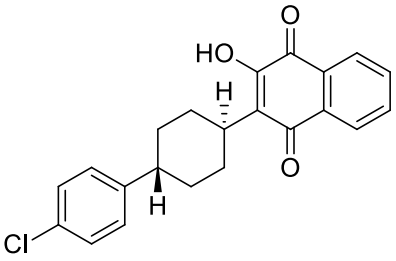
pyrimethamine emerged and thus the antibiotic sulfadoxine (Scheme 1.1), was combined with pyrimethamine doses to provide a synergistic effect.[18] The pyrimethamine-sulfadoxine compounds target two enzymes, namely, dihydrofolate reductase and dihydropteroate synthetase which are present in the folate biosynthesis pathway of the parasite.[19] A study in 1996 showed that the combination of atovaquone (Scheme 1.1), an inhibitior of the mitochondrial electron transport chain in the parasite, and either tetracycline (an antibiotic that inhibits protein synthesis) or proguanil showed cure rates of greater than 95%. Atavaquone cannot be given as the sole therapy as this can lead to recurrence of parasitemia.[16,20–22]



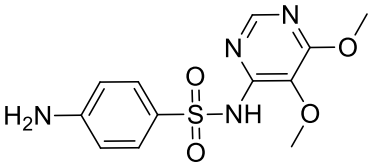
Proguanil



Pyrimethamine



Atovaquone



Sulfadoxine

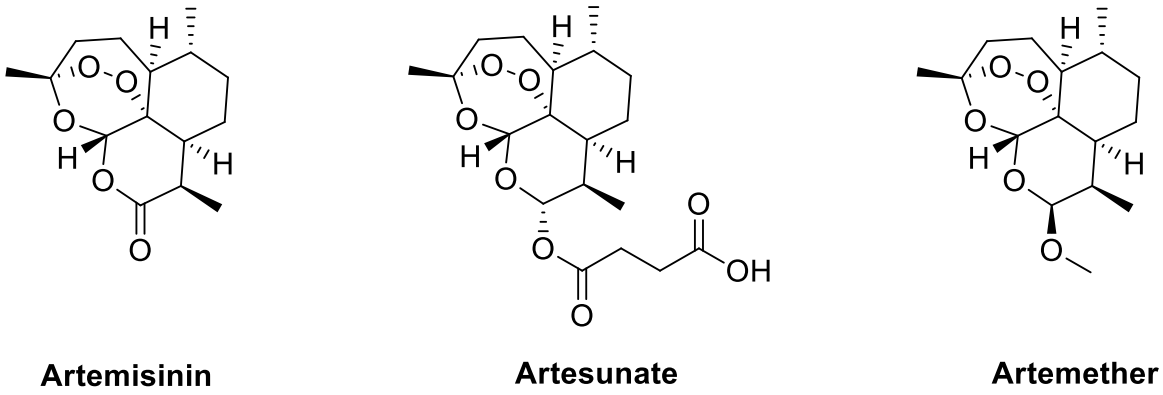
Scheme 1.1: The pyrimidine based antifolate drugs proguanil, pyrimethamine and sulfadoxine, and the antiparasitic atovaquone.

1.5.2 Artemisinin Derivatives

Due to significant resistance shown to any single drug treatment, the recommended treatment for malaria is ACTs.[7,8] ACTs combine artemisinin and a partner drug, where artemisinin has a fast mode of action that kills gametocytes and alleviates symptoms, while the longer lasting partner drug kills the rest of the parasites after the effect of artemisinin has worn off. This prevents susceptibility to only one antimalarial agent.[8,16,23–26] Artemisinin (Scheme 1.2) is a sesquiterpene lactone that contains an endoperoxide bridge and was first extracted from the leaves of the Chinese wormwood, also known as *Artemisia annua*, in the

late 1970s to early 1980s by Tu Youyou and co-workers, which was recognised by the award of a Nobel Prize in Physiology and Medicine in 2015.[9,26,27] Before this, the leaves of *Artemisia annua* had been used to treat chills and fevers for over 2000 years.[26] Several mechanisms of action for this antimalarial have been debated throughout the years however, these have all been shown to act during the blood stage of the parasite.[28,23] It has been suggested that cleavage of the endoperoxide bridge is catalysed by Fe(II) producing radicals that destroy the parasite.[28,29] Two mechanisms have been proposed, namely, the reductive scisson model and the open peroxide model. The reductive scisson model involves Fe(II) binding to artemisinin and after electron transfer takes place it produces oxygen radicals that rearrange to give carbon radicals which are believed to alkylate malaria proteins. The open peroxide model proposes that ring opening is either driven by protonation in the acidic food vacoule, assisted by oxygen or complexation with Fe(II). Hydroperoxide is formed from heterolytic cleavage followed by addition of water. This hydroperoxide residue is capable of oxidising proteins.[28] Also Fenton[30] degradation of the hydroperoxide produces oxygen radicals that are toxic to the parasite.[28]

The synthesis of derivatives of artemisinin, namely, artemether, the methyl-ether derivative, (Scheme 1.2) and artesunate, the succinic acid analogue, (Scheme 1.2) were carried out in the 1980s. However, all these compounds are broken down into the same metabolite, dihydroartemisinin, and thus these compounds are pro-drugs. [31] It has been noted that monotherapy with artemisinin compounds results in recurrence of parasitemia after treatment.[26] This is mitigated with the use of ACTs and according to the WHO 409 million doses of ACTs where given out in 2016.

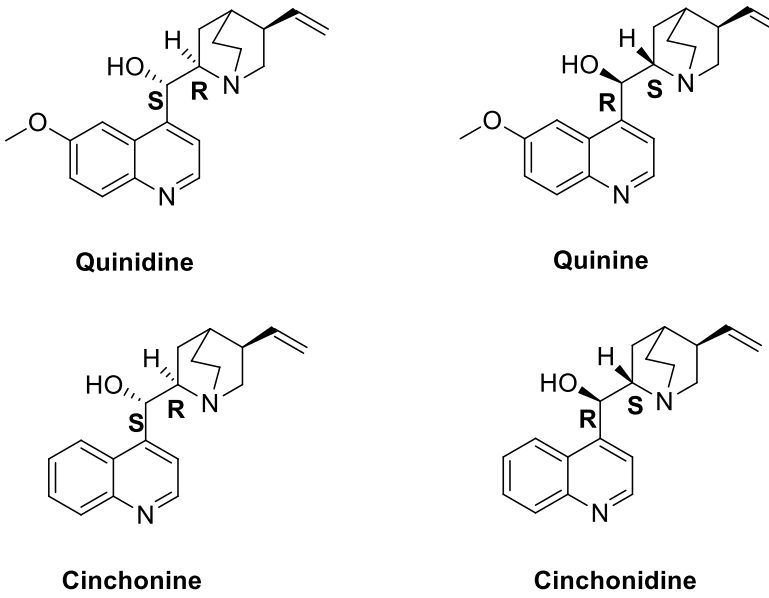


Scheme 1.2: The artemisinin derivatives.

1.5.3 Quinolines and Derivatives

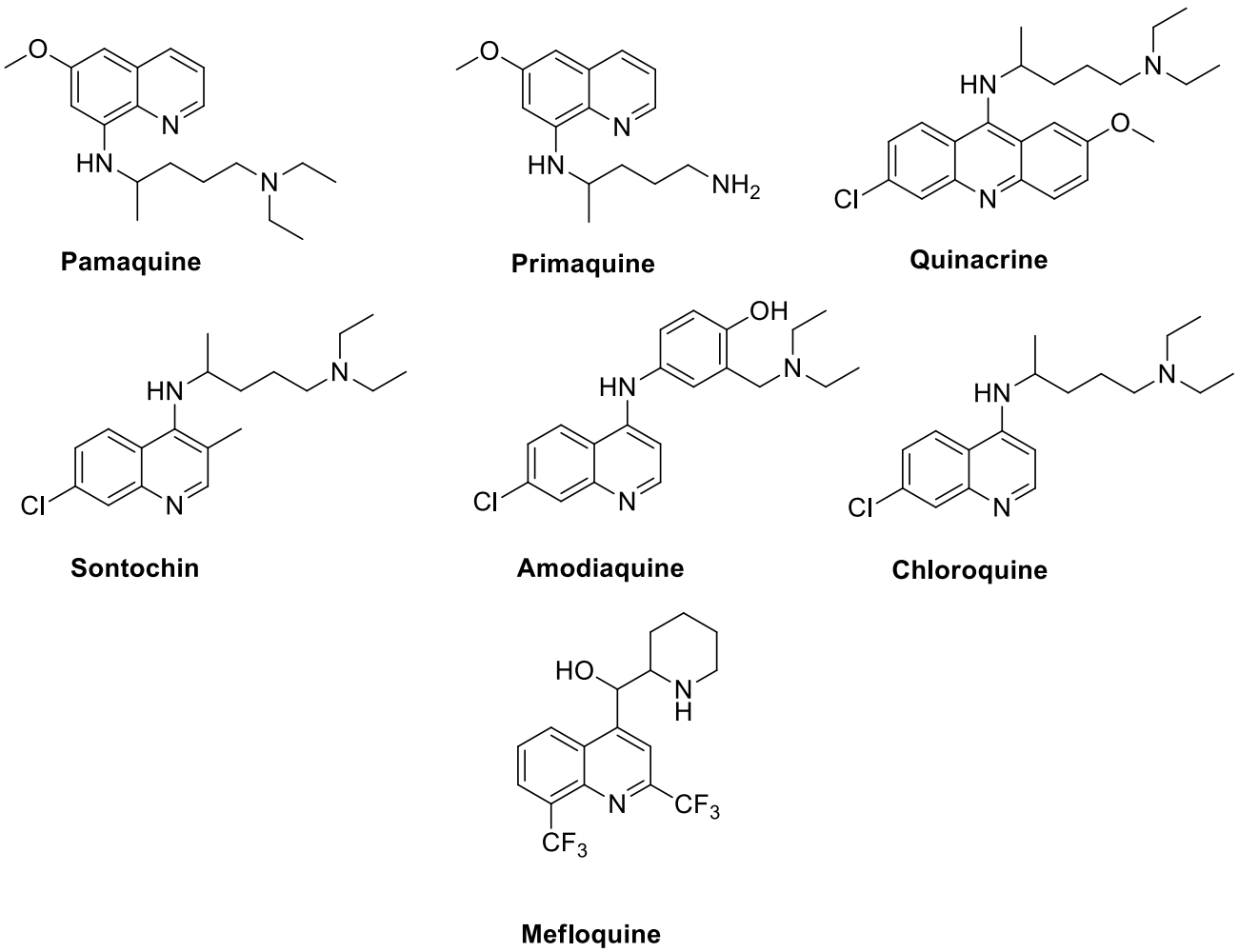
The first malaria chemotherapy, obtained from the crushed bark of the *Cinchona* tree was discovered in 1630 by Spanish Jesuits who imported it into Europe from Peru. The bark of the tree was given the latin name *Cinchona* after a fictional tale of a countess of Cinchon who was cured of malaria using this crushed bark. The first attempt at isolation of the compounds found in this bark was in Portugal, by Gomes in 1810. After 10 years, the alkaloids, quinine and cinchonine, and later their diastereomers quinidine and cinchonidine were finally isolated by Caventou and Pelletier.[6,32,33] The drug's popularity led to the destruction of the Cinchona forests in Peru, therefore the Dutch and British smuggled seeds out of Peru and started their own plantations in Java and India. From these plantations roughly 10 million kg of quinine was produced annually.[6] The structure of quinidine, one of the active compounds found in the tree bark was finally elucidated in 1908.[32]

Cinchonidine and cinchonine are diastereomers and upon substitution of the hydrogen on the 6-position of the the quinoline ring with a methoxy group, quinidine and quinine are formed. The four chiral alkaloid compounds isolated from the tree are shown in Scheme 1.3 with their respective stereochemistry.[34]



Scheme 1.3: The structures of the alkaloids isolated from the Cinchona bark and their relative stereochemistry.

During the first World War, the Germans found themselves cut-off from the world supply of quinine and thus synthetic analogues needed to be discovered. The starting point of this was the observation by Ehrlich that methylene blue showed some effect on malaria patients and thus in the 1920s, Schönhofer and Wingler investigated thiazine derivatives. This led to the discovery of the first 8-aminoquinoline, pamaquine[35] (Scheme 1.4), which showed only activity against avian plasmodia. Primaquine (Scheme 1.4) was synthesized as an active, less toxic analogue. German scientists replaced the quinoline ring by acridine while also adding a basic side chain. Atebrin, now called quinacrine (Scheme 1.4) was found to be most active against the asexual forms of *P. falciparum*. German scientists continued with their research and found that attachment of basic side chains to the 4-position of the quinoline moiety brought about increased activity. Following this, Andersag synthesized sontochin[36] (Scheme 1.4) and resochin in 1934 but, initial testing revealed resochin to be toxic and it was thus abandoned. This error was rectified 8 years later when a screen of more than 17 000 compounds against avian plasmodia was conducted in the USA in 1943 in which resochin (now called chloroquine) and amodiaquine (Scheme 1.4) were shown to have increased theraputic activity compared to sontochin and were found to have acceptable toxicity. Clinical trials of chloroquine started in 1943, after which it was the antimalarial of choice for 25 years.[6,10] After the rediscovery of chloroquine, in 1944 the first synthetic pathway that could not be used for commerical use for quinine, was described by Woodward and Van Doering.[6] This took 36 years after the discovery of the structure to be synthesized. Due to problems with chloroquine resistance, another large screen of over 300 000 compounds was carried out with the 4-quinolinemethanols emerging as the most promising group.[6,10] These compounds were analogues of quinine but had a strong photosensitizing action. This was rectified with the synthesis of mefloquine (Scheme 1.4) which has proven useful against chloroquine-resistant parasites.[37]



Scheme 1.4: The most influential antimalarial drugs synthesized from the 1920s till 1984.

Throughout the years many newer antimalarial drugs have been synthesized targeting different stages of the life cycle of the parasite as well as different organelles in the parasite. Figure 1.4 shows a summary of all the antimalarial drugs currently used and their targets in the malaria parasite.

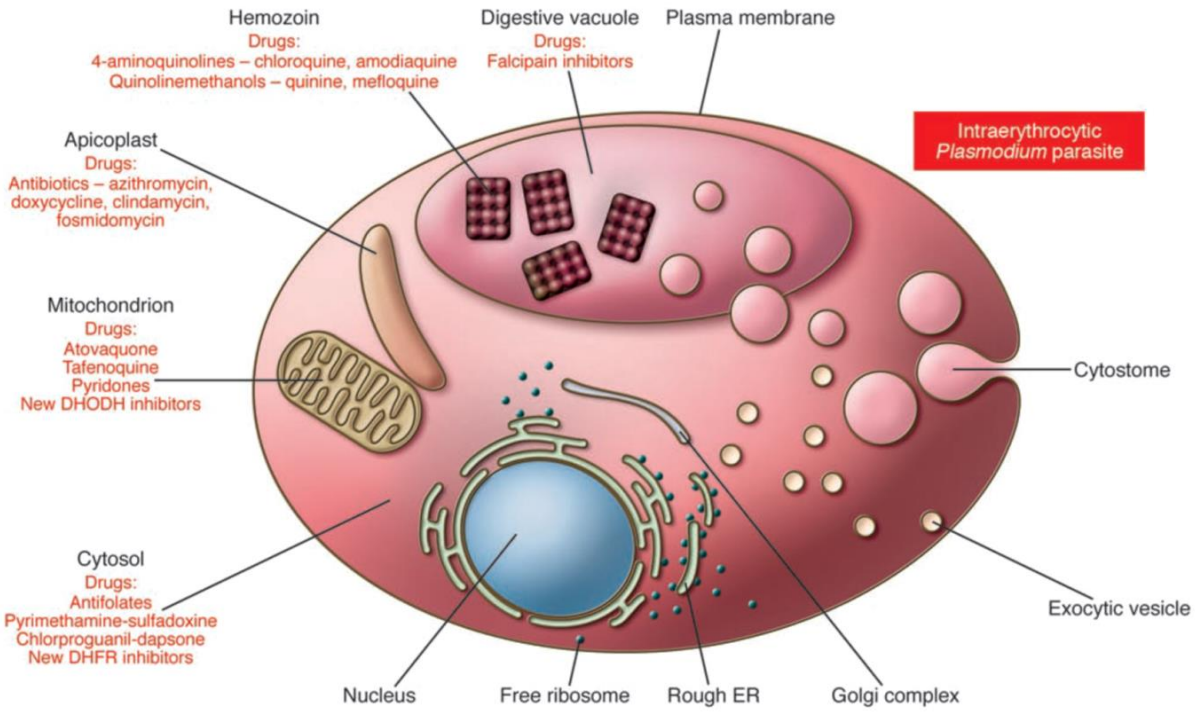


Figure 1.4: A summary of antimalarial drugs and their respective targets. Reprinted with permission from Greenwood *et al*.[8] Copyright © 2008 American Society for Clinical Investigation.

1.6 Resistance to Known Drugs

According to the WHO the malaria death rate has decreased by between 27 and 44% depending on the region between the years 2010 and 2016 mostly due to the use of ACTs.[7] However, resistance to known antimalarial drugs has had a devastating effect on treatment of malaria. There is a correlation between extended use of a drug and the development of resistance. For example in 1961, 84 000 tons of chloroquine was supplied for inclusion in salt. This quickly brought about resistance in those areas where this salt was used.[6]

However, the first recorded case of *P. falciparum* chloroquine resistance was in Thailand in 1957 when standardised tests were introduced to assess the *in vivo* response to drugs in *P. falciparum*.[38] After this, an attempt was made to eradicate the world of malaria by using chloroquine and insecticides as treatments. This pressure for the parasite to survive brought about rapid spread of resistance.[32] However, chloroquine still remains the main treatment for individuals infected with *P. vivax*.[7] Figure 1.5 shows the global distribution of chloroquine

resistance in 2002. This shows that even in the absence of any drug pressure after 45 years, the chloroquine resistant strains of *P. falciparum* still persist.[33]

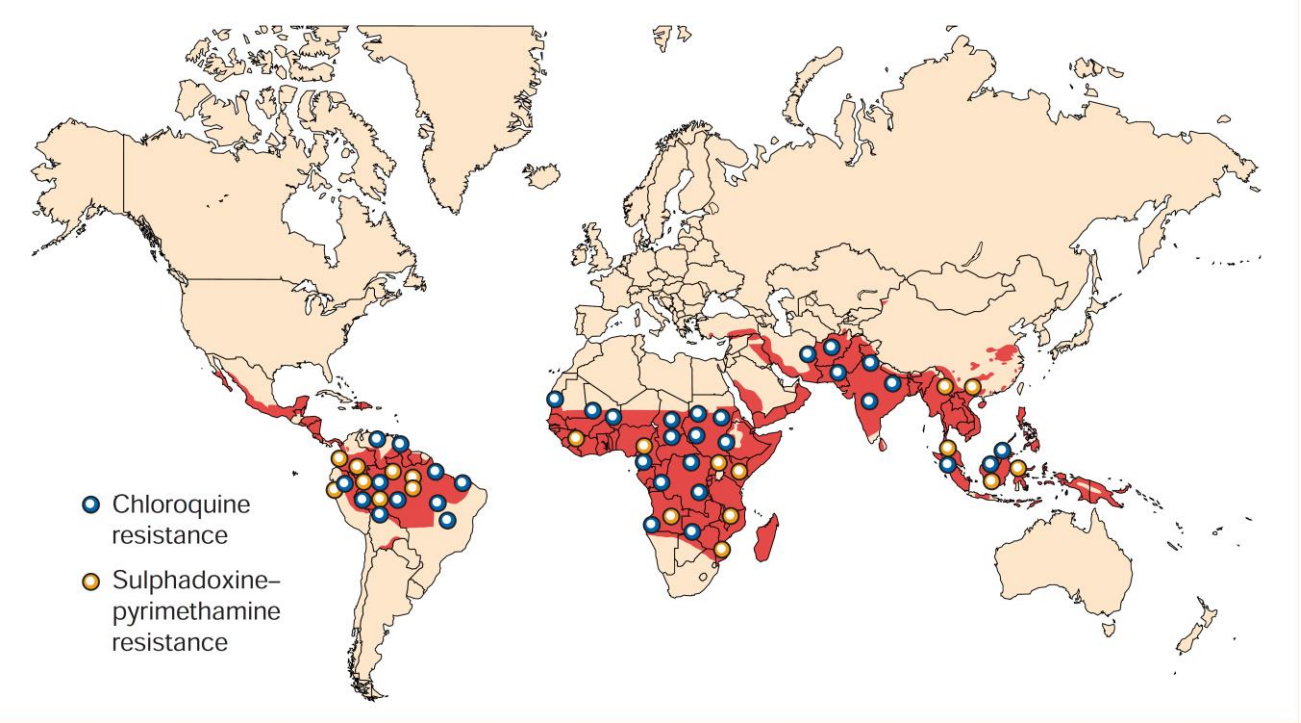


Figure 1.5: The global status of chloroquine and sulphadoxine-pyrimethamine resistance in 2002 accoding to the WHO. Reprinted with permission from Ridley. [33] Copyright © 2002 Nature.

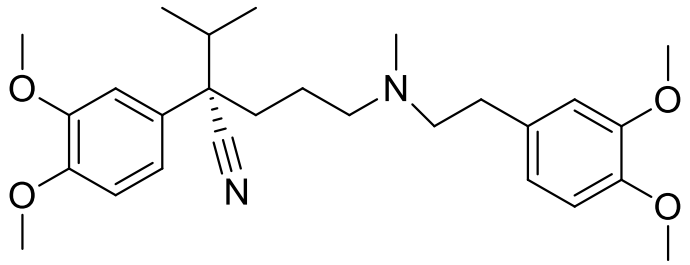
Even though quinine is the oldest antimalarial known to man, resistance has been slow to build against this drug since it is used much less frequently.[6,39] Contributing factors to this is the bitter taste, side-effects such as hypoglycaemia and the dosage of three times daily for seven days.[16] Resistance to the antifolate drugs, pyrimethamine and sulphadoxine, spread rapidly due to single point mutations in the dihydrofolate reductase and dihydropteroate synthase which lower the binding affinity of these drugs to the enzymes.[40–43] Similary, atovaquone resistance results from point mutations in the gene *cytB*, which encodes cytochrome b.[41,42] Mefloquine resistance has also developed rapidly due to overuse in Thailand, however it is still used in combination with artemisinin. Mefloquine resistance has been linked to the *P. falciparum* multidrug resistance gene (*Pf*mdr1). However, similar effects have been shown with halofantrine and quinine and thus this is most likely not the only cause of resistance to mefloquine.[6,9,41–44] There is usually very little evidence of cross-resistance

between the 4-aminoquinolines and quinoline methanol compounds and thus compounds such as mefloquine are used on chloroquine resistant strains of *P. falciparum*.[6,44]

Until 2008, resistance to artemisinin had not been reported. The first evidence was found in Cambodia, where the clearance rates were found to be more than double compared to the median expected for artemisinin. This was due to monotherapy with artemisinin carried out in this region despite literature indicating this should never be done. This resistance is due to the short half-life of the drug and the WHO recommends combination with drugs that have longer half-lives to mitigate this problem such as mefloquine.[7,45]

By far the most researched mechanism of resistance is that of chloroquine due to its previous potency as an antimalarial drug. The main observation was that chloroquine accumulates in the digestive vacoule in high concentrations. This is substantially reduced in the chloroquine resistant strains of *P. falciparum*. This key finding sparked much debate about the mechanism of this phenomenon. However, through many years of research the actual mechanism of resistance was discovered.[6,46,47]

The initial proposed mechanism was that of drug efflux out of the digestive vacuole in the resistant strains. A similar mechanism had been shown with different types of cancer cells and it was also shown that this drug resistance can be reversed with exposure to verapamil (Scheme 1.5) similar to those of drug resistant cancer cells.[48,49] Several other mechanisms were also proposed including, increased detoxification of the chloroquine-haematin complex, altering chloroquine uptake by the digestive vacuole membrane or H^+ flux and prevention of the chloroquine-haematin complex from forming.[49–51]



Scheme 1.5: The chemical structure of the chloroquine resistance reversal agent, verapamil.

All of these mechanisms were proved incorrect when Fidock *et al* [52] discovered a digestive vacuole transmembrane protein, aptly called *P. falciparum* Chloroquine Resistance Transporter (*Pf*CRT), which was the primary reason for resistance. The study in 2000 showed

a gene containing eight point mutations in the chloroquine resistant (Dd2) strain compared to that of the chloroquine sensitive (HB3) strain of *P. falciparum*. Upon exchange of the alleles of the resistant to the sensitive strains, lowering of activity of chloroquine was demonstrated with verapamil showing reversal of resistance as with the chloroquine-resistant strain. It was further observed that the most crucial mutation for resistance was the K76T mutation, an exchange of a lysine by a threonine amino acid at position 76.[52,53] Another study by Roepe showed binding of *Pf*CRT to a fluorescently tagged chloroquine analogue using spinning disk confocal microscopy.[48,54] Juge *et al* in 2015 was able to demonstrate active transport of chloroquine by *Pf*CRT across a proteoliposome membrane from lower pH to that of higher pH comparable to that of a digestive vacuole. It was also shown that verapamil was able to pass through the *Pf*CRT liposome assembly and inhibit chloroquine resistance.[55]

1.7 Mechanism of Action of the Quinoline Based Compounds

Initial studies by Parker and Irvin demonstrated that chloroquine can bind DNA and thus it was proposed that it can inhibit DNA replication and RNA synthesis in the malaria parasite.[56–58] This, however does not explain the fact that chloroquine kills *P. falciparum* at concentrations three times lower than those needed to be cytotoxic.[6] A mechanistic clue came in 1970 when it was observed that chloroquine only acts against the blood stage of the parasite. This led to the belief that chloroquine interferes with the feeding process of *P. falciparum*. It was only in 1986 that Fitch *et al* [59] proposed that chloroquine interferes with haemozoin formation by complexing with haem. Later, Slater and Cerami complemented this argument in 1992 by demonstrating inhibition of β-haematin formation *in vitro* using chloroquine.[60] It has been shown that haem concentrations in the digestive vacuole can be anywhere between 200-500 mM if not converted to non-toxic haemozoin by the parasite (Section 1.4). Thus this is an ideal target for malaria chemotherapy.[6]

In order to attempt to verify compounds as haemozoin inhibitors before techniques for directly observing this had been developed, significant correlations between the haemozoin inhibition and parasite activity were shown. The first of these was demonstrated by Dorn *et al* [61–63] where a significant correlation between quinoline antimalarials β-haematin inhibition activity and NF54 *P. falciparum* parasite activity was shown (Figure 1.6). The assay was carried out using radio-labelled [^{14}C]-haemin with eight known antimalarial drugs, namely,

chloroquine, halofantrine, quinacrine, amodiaquine, pyronaridine, quinine, mefloquine and primaquine as well as an experimental bisquinoline compound named Ro 48-6910.[61]

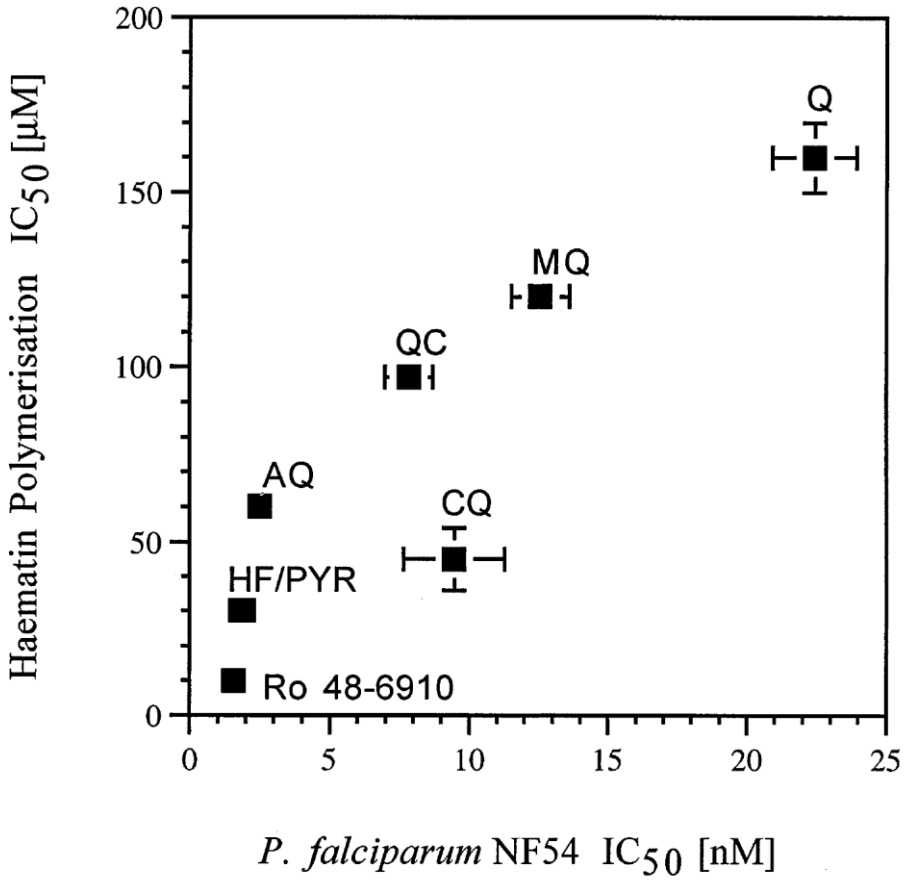


Figure 1.6: The correlation of β-haematin inhibition (called haematin polymerisation) versus the NF54 *P. falciparum* parasite growth inhibtion activity of eight compounds namely, quinine (Q), mefloquine (MQ), quinacrine (QC), amodiaquine (AQ), chloroquine (CQ), halofantrine (HF), pyronaridine (PYR) and bisquinoline Ro 48-6910. Reprinted with permission from Dorn *et al*.[61] Copyright © 1998 Elsevier Science.

Another study in 1999, by Vippagunta *et al* [64] related the β-haematin inhibition values of thirteen chloroquine analogues to both the NF54 (chloroquine sensitive) and K1 (chloroquine resistant) parasite growth inhibition values (Figure 1.7). However, in this study the β-haematin inhibition values were normalised using the association constants of the drug with haem improving the correlation substantially. This showed the importance of the strength of binding of drugs to haem and how it influences parasite activity.[64]

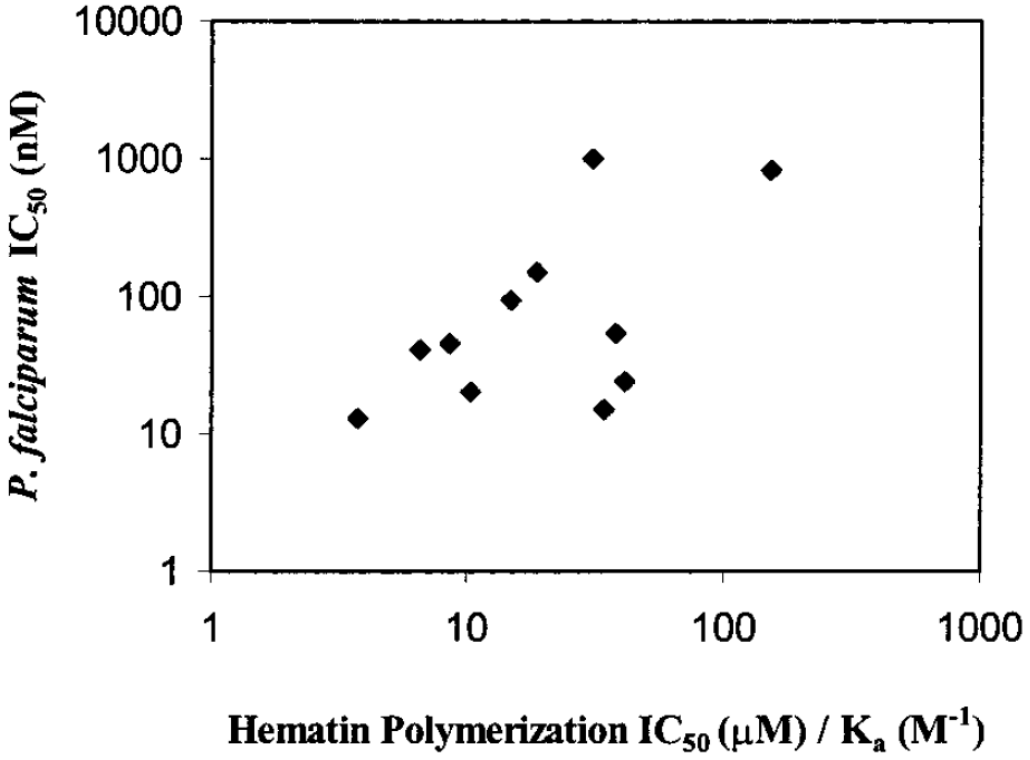


Figure 1.7: The correlation of normalised β-haematin inhibition values using association constants of the drug with haem versus the parasite growth inhibition values with three compounds omitted due to lack of β-haematin inhibition activity. Reprinted with permission from Vippagunta *et al*.[64] Copyright © 1994 American Chemical Society.

Kaschula *et al* [65] also reported a correlation using chloroquine derivatives in 2002. Using an assay based upon differential solubility of haematin in dimethyl sulfoxide (DMSO) to calculate β-haematin inhibition values,[66] no direct correlation was found. However, when the vacuolar accumulation ratio (α) was multiplied by the log(*P.f* IC_{50}) a trend emerged (Figure 1.8). α was based on the pKa values of each compound and this ratio was divided by the α of chloroquine to determine the relative extent of pH trapping within the digestive vacoule of each compound seperately.[65] Development of the pyridine ferrichrome method by Ncokazi and Egan [67] in 2005, which was found to be a more reliable method to measure β-haematin inhibition, was used to re-examine the activity of these compounds. Upon multiple correlation analysis using the vacuolar accumulation ratio and β-haematin inhibition values, a similar trend to that of the orignal Kaschula *et al* data was shown (Figure 1.9). There were similar correlations reported using non-quinoline compounds.[68,69]

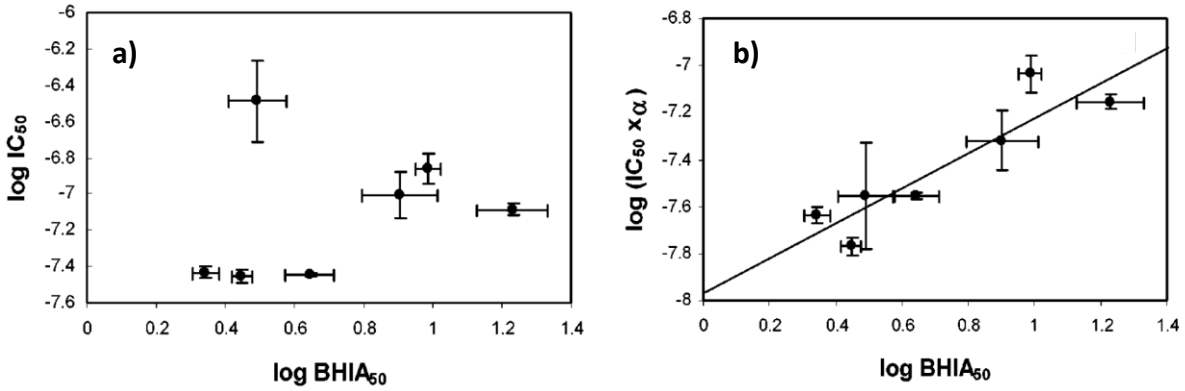


Figure 1.8: a) Structure-activity relationship (SAR) showing no correlation and b) correlation between log(*P.f* IC_{50} (α)) and the D10 parasite growth inhibition activity of chloroquine analogues. Reprinted with permission from Kaschula *et al*.[65] Copyright © 2002 American Chemical Society.

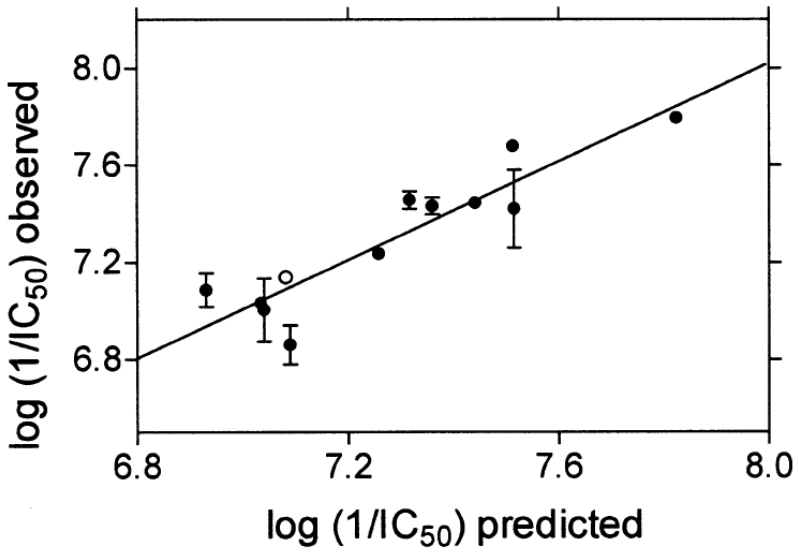


Figure 1.9: Multiple correlation analysis with the equation of log(1/*P.f.* IC_{50}) = 0.61(log[1/IC_{50}(Phiβ)]) + 0.17(logVAR) + 6.98 where VAR is vacuolar accumulation ratio. Reprinted with permission from Ncokazi and Egan. [67] Copyright © 2005 Elsevier.

In 2006, Roepe *et al* [70] reported using spinning disk confocal microscopy to observe live parasites in red blood cells. These cells were dosed with chloroquine and observed over 37 hours by measuring pixel density in z-stack (z-slice) images which were reconstructed to form a live image of the parasite in real time. It was shown that chloroquine inhibits haemozoin formation from 22 to 29 hours depending on the dose for both chloroquine sensitive and

resistant strains (Figure 1.10). Furthermore it was observed that the use of verapamil lowered the IC_{50} of the chloroquine resistant strain, Dd2.

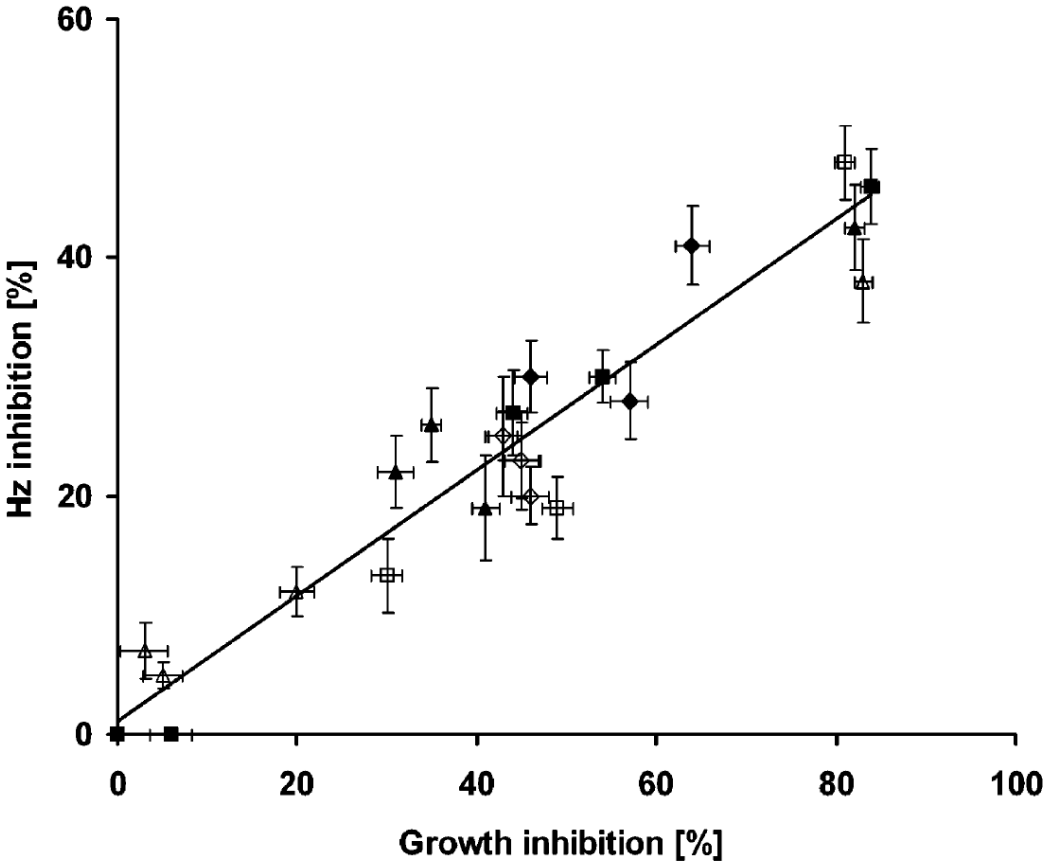


Figure 1.10: The correlation of percent haemozoin inhibition versus percentage parasite growth inhibition between 25 and 30 h for chloroquine sensitive (HB3) strain (empty markers) with verapamil (triangle), without verapamil (squares), combination of verapamil and chloroquine (diamond) and the chloroquine resistant strain (Dd2) with filled markers. Reprinted with permission from Roepe *et al*.[70] Copyright © 2006 American Chemical Society.

From this study it is clear that decreased haemozoin levels in live *P. falciparum* cells support the haem detoxification pathway via haemozoin inhibition. However, this cannot validate haemozoin as the target as this reduction of haemozoin formation may be due to other factors such as reduced haemoglobin uptake. Therefore, Combrinck *et al* [71] devised the cellular haem fractionation assay in which the amount of free haem and haemozoin can be measured in cultured *P. falciparum* after exposure to known quinoline antimalarial drugs. Chloroquine showed the largest dose-dependant decrease in haemozoin and increase in free haem of all the antimalarial drugs tested with the parasite-survival curve showing the 50%

inhibitory concentration (IC_{50}) of the drug at the cross-over point (Figure 1.11). It was also observed that there were no changes in the levels of free haem and haemozoin when dosed with 2.5 times the IC_{50} of the antifolate drugs, pyrimethamine and the pyrimethamine-sulfadoxine combination. This was expected because these drugs are known to have a different mode of action. This study shows the most evidence in support of the haemozoin inhibition mechanism proposed by Fitch to date.

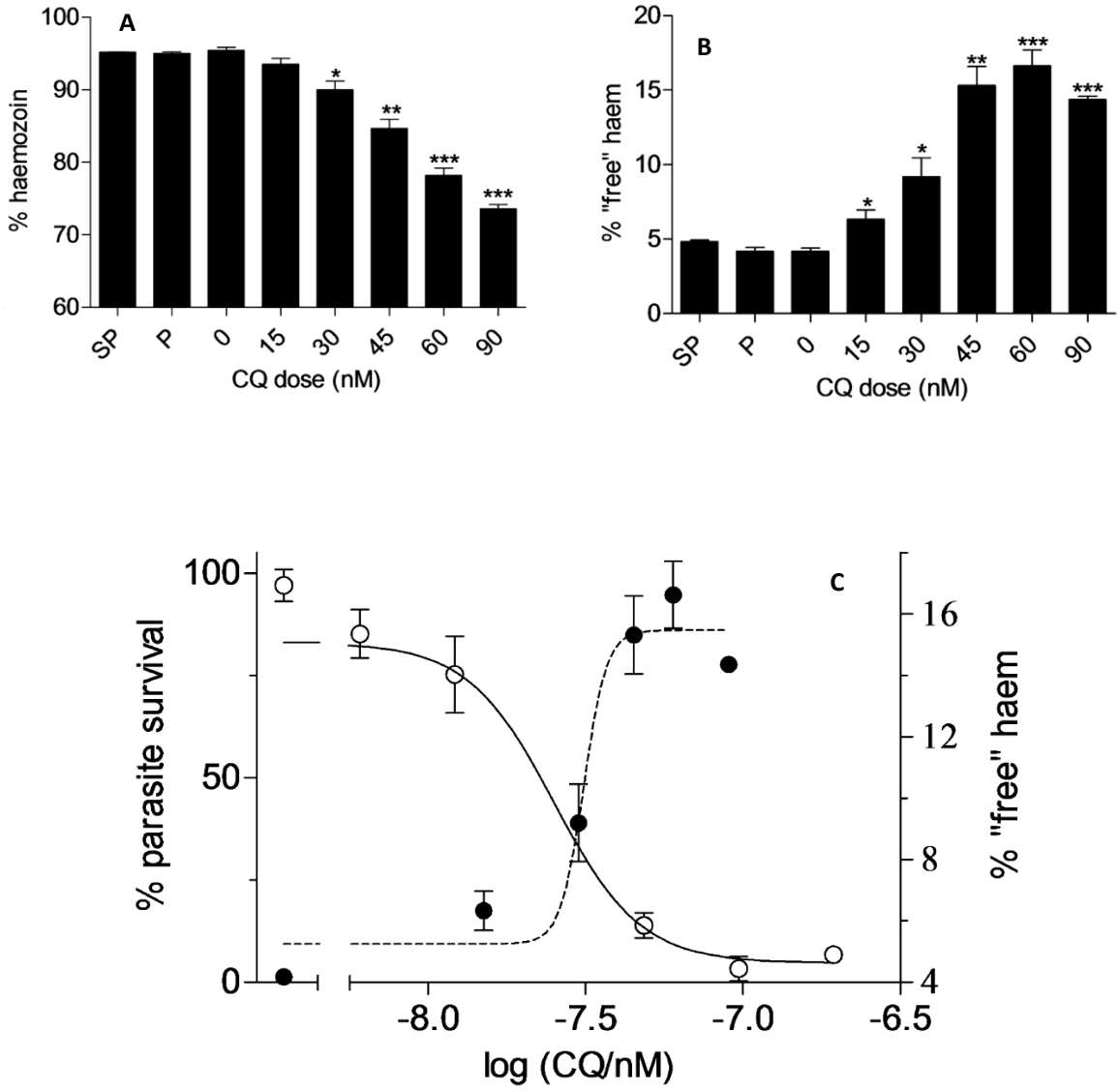


Figure 1.11: Graphs showing the effect of chloroquine dosing in *P. falciparum* on a) % haemozoin, b) % free haem and c) the % parasite survival with the % free haem. The negative control is pyrimethamine/sulfadoxine (SP). Reprinted with permission from Combrinck *et al.*[71] Copyright © 2012 American Chemical Society.

1.8 β-haematin and Haemozoin

With the knowledge that quinoline based compounds as well as other antimalarials inhibit haemozoin formation, it is appropriate to discuss the process by which haemozoin is formed. Even though this has been studied for decades the exact mechanism of haemozoin formation is not known.

1.8.1 Discovery of the Structure of Haemozoin and its Synthetic Analogue, β-haematin

The first report of haemozoin was in Italy by the renowned physician Lancisi in 1717.[12,15] Later in 1897, Surgeon-Major Ross also noticed peculiar pigmented cells.[4] Only in 1911 did W. H. Brown identify this pigment to consist of haem after it was originaly believed to be melanin.[72] The first isolation of haemozoin was demonstrated by Fitch *et al* in 1987 where the erythocytes were lysed and washed. Remaining proteins and lipids were broken down by proteases and the mixture was extracted using methanol and chloroform to yield the haemozoin crystals.[59] In 1991, Slater *et al* [73] characterized haemozoin by using various spectroscopic techniques and comparing it to synthesized haemozoin, also known as β-haematin. An absorbance peak at 650 nm was observed in the UV-visible absorbance spectrum, characteristic of aqueous haematin which was already observed by Fitch *et al* [59] in 1987. This was further confirmed with the elemental composition of haematin and the extracted haemozoin showing the same data. Upon dissolution of the extracted haemozoin in 0.1 M NaOH, the UV-visible absorbance spectrum as well as the mass spectrum were found to be identical to that of haematin further correlating the structure of haem to that of haemozoin. Haemozoin was further analysed using Fourier-transform infrared (FT-IR) spectroscopy where it was shown that this spectrum differed from that of haem containing two additional bands at 1664 and 1211 cm^{-1} which are characteristic of C=O and C-O stretching frequencies. Its was mentioned that this might arise from direct coordination of a carboxylate group of one molecule to the iron centre of another which was previously observed in an extensive series of Fe(III) monodentate complexes with coordination by acetate which contained a similar infrared (IR) peak at 1660 cm^{-1}. This coordination of the propionate to the iron centre was only seen in the spectrum of haemozoin and not haematin. Using electron spin resonance (ESR) spectroscopy it was claimed that the ferric iron was in the low-spin state. However, this was proved incorrect later by Bohle *et al* [74] using Mossbauer

and variable-temperature electron paramagnetic resonace (EPR) spectroscopy where in fact this ferric iron centre is in the high-spin state. A year ealier, Bohle *et al* [75] also confirmed using powder X-ray diffraction (PXRD) patterns that haemozoin and β-haematin are identical. This data suggested that haemozoin takes the form of polymer chains hydrogen bonded by the propionate groups shown in Figure 1.12.

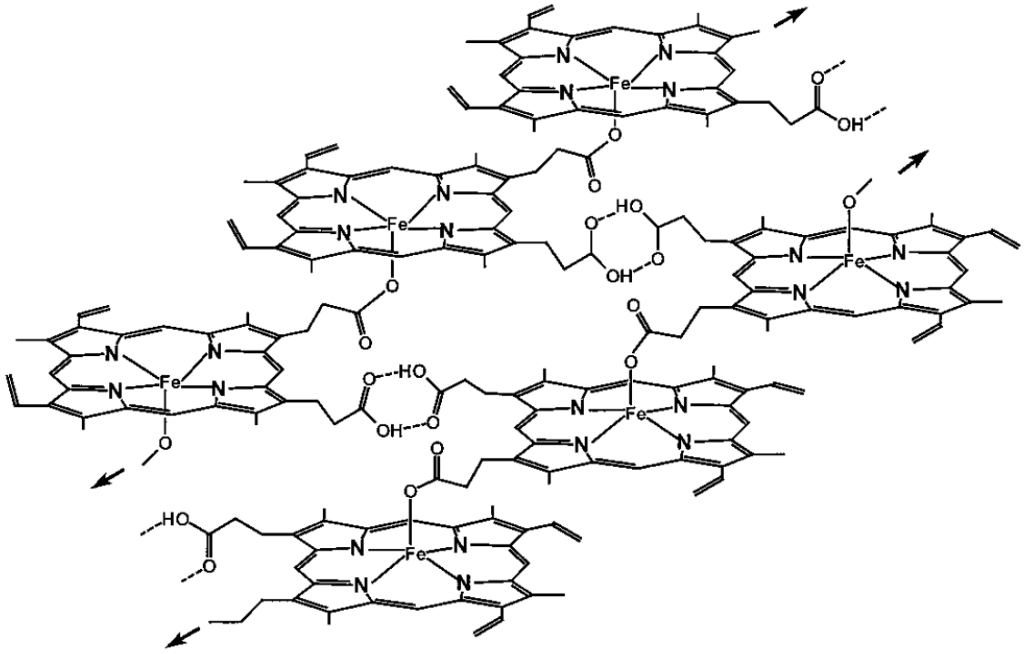


Figure 1.12: The Incorrect polymer structure of haemozoin proposed by Bohle *et al* where the arrows indicate direction of polymer elongation. Reprinted with permission from Bohle *et al*. [75] Copyright © 1997 The American Society for Biochemistry and Molecular Biology.

Later, in 2000, Pagola *et al* [76] corrected this structure by solving the crystal structure using PXRD patterns obtained from synchrotron radiation by Rietvelt refinement. The structure of β-haematin was observed as dimers that interact via hydrogen bonding between propionate groups. These dimers then pack togeather via π-π interactions forming a crystal structure (Figure 1.13).

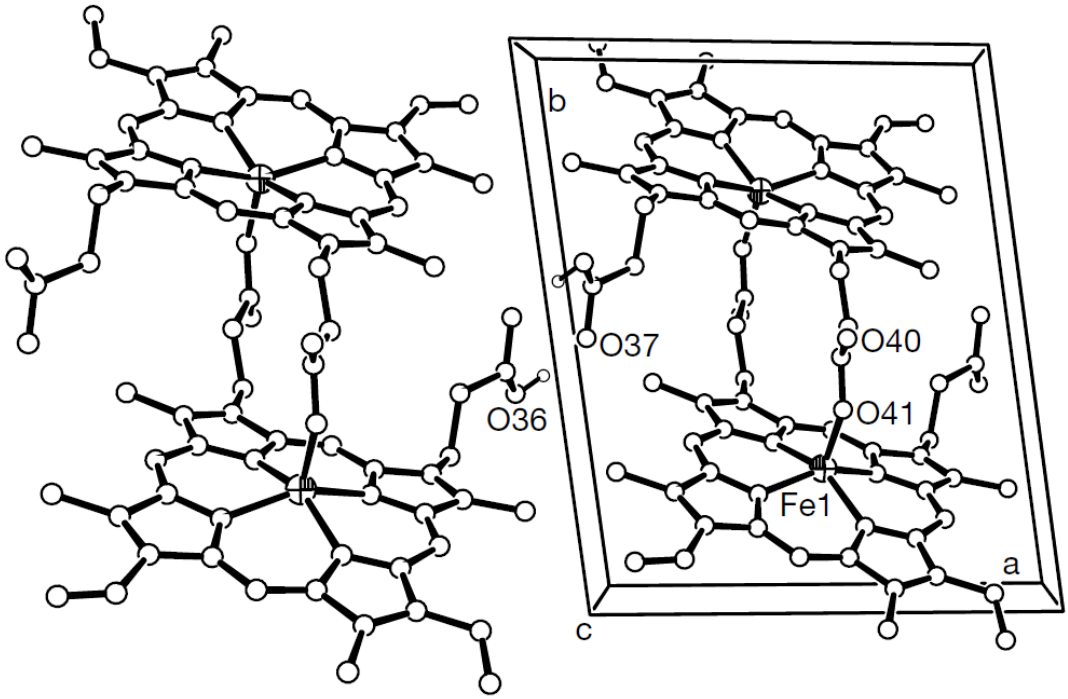


Figure 1.13: The actual crystal structure of two β-haematin dimers hydrogen bonding via the propionate groups. Reprinted with permission from Pagola *et al*.[76] Copyright © 2000 Nature.

1.8.2 Haemozoin Formation Mechanism

There have been many hypothesized mechanisms proposed that attempt to explain the formation of haemozoin including, protein mediated formation[77,78], enzyme catalysed[60], spontaneous formation[79] and lipid mediated[62,63]. Of these the most evidence points towards the lipid mediated mechanism whereby several studies have shown succesful formation of β-haematin in the presence of lipids.[80,81] Haemozoin crystals have also been observed via microscopy and found to be in close proximity to lipids in *P. falciparum*.[81] Recently, Kuter *et al* [82] have demonstrated that glycerolipids are responsible for the formation of haemozoin rather than phospholipids. The most abundant of these were found to be diacylglycerols rather than monoacylglycerols which was identified by using thin layer chromatography (TLC) analysis. With the help of extensive computational studies involving this lipid and haem as well as experimental findings a mechanism describing the intial stages of haemozoin formation was proposed (Figure 1.14). The main steps involve haem partitioning between lipid and aqueous phases after which haem is aggregated within the lipid, iron coordination occurs with the propionate groups of other haem molecules and finally crystal growth occurs.[82]

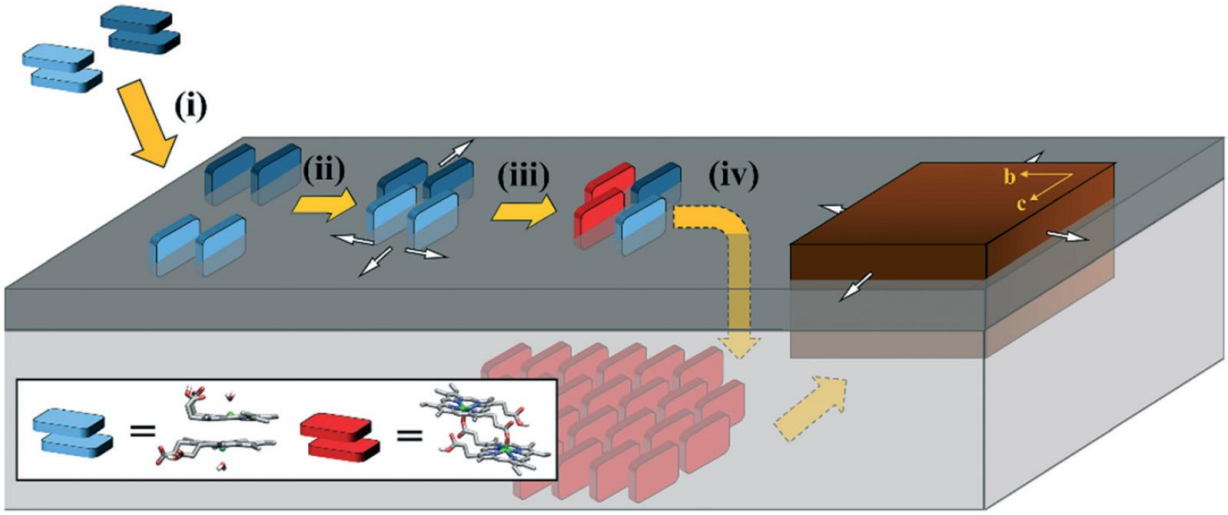


Figure 1.14: The proposed mechanism of haemozoin formation using the water-lipid interface where the dark grey band represents lipid headgroups while the light grey band represents lipid tails. Key steps are i) partitioning of haem into lipid, ii) aggregation of haem in lipid, iii) iron coordination and iv) crystal growth. Reprinted with permission from Kuter *et al*.[82] Copyright © 2016 the Royal Society of Chemistry.

1.8.3 Haemozoin Inhibition Mechanism

The mechanism of haemozoin inhibition is not certain even though it is known that haemozoin-inhibiting drugs cause a decrease in haemozoin and increase in free haem in the parasite (Section 1.7). Dorn *et al* [62] demonstrated that quinoline antimalarials bind to the μ-oxo dimer thereby reducing the amount of monomeric haem available for formation of haemozoin. This presence of μ-oxo dimer formation was confirmed by Kuter *et al* [83] who further observed a complex of chloroquine and Fe(III)PPIX with a stoichiometry of 2:4 using magnetic circular dichroism, UV-visible absorption, diffusion techniques and magnetic susceptibility. However, it is not certain if the complex is responsible for haemozoin formation inhibition. Some mechanistic suggestions indicate that this complex can bind to haemozoin and disrupt or terminate growth.[77] Another suggestion is that the antimalarial drug in question could bind to the fastest growing crystal face to inhibit further growth of the crystal. Further computational studies by Buller *et al* [84] based on the crystal structure of Pagola *et al* [76] have shown chloroquine manually docking on the fastest growing face of the β-haematin crystal, (001), as shown in Figure 1.15. Complementing this work Olafson *et al* [80] have shown

four different adsorption sites on the β-haematin crystal but, it was also shown that chloroquine adsorbs preferably to the step sites on the (001) crystal face.

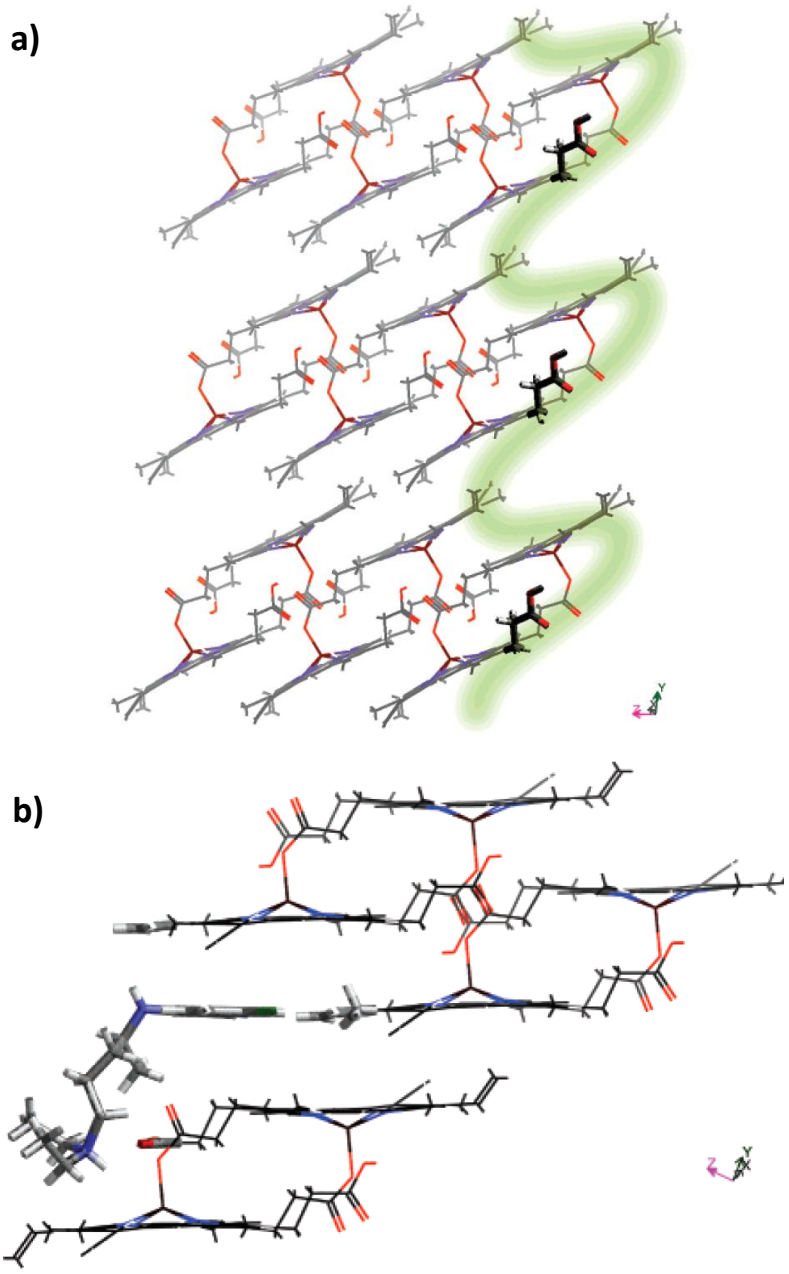


Figure 1.15: a) The (001) face of the β-haematin crystal structure with crevices shown in green and b) chloroquine interacting with the (001) face of the β-haematin crystal structure. Reprinted with permission from Buller *et al*.[84] Copyright © 2002 American Chemical Society.

It should be noted that even though haemozoin inhibitors follow the same mode of action, resistance arises differently to each scaffold and thus cross-resistance does not occur

between the quinolines and other scaffolds. It is therefore still viable to still consider haemozoin inhibtion as a drug target.

1.8.4 Haem-drug Complexes

The first haem-drug complex crystal structure of halofantrine and haem was reported in 2008 by de Villiers *et al*.[85] This crystal structure (Figure 1.16) revealed the deprotonated hydroxyl group of halofantrine coordinating to the iron (III) centre and π-π stacking interactions between the porphyrin ring system and that of the phenanthrene ring. Using the UV-visible absorbance data of this complex it was shown that it was similar to that of the haem-quinidine and haem-quinine complexes. Consequently, similar structures were proposed. This was later confirmed by de Villiers *et al* [86] when the actual crystal structures were reported (Figure 1.16).

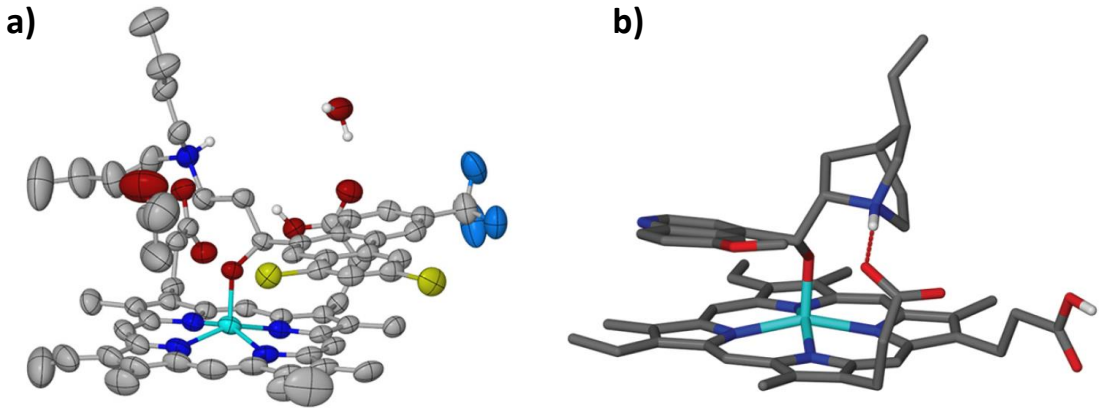


Figure 1.16: Crystal structures of a) the halofantrine-haem complex and b) the quinine-haem complex. Reprinted with permission from de Villiers *et al* 2008 [85] and 2012 [86] respectively. Copyright © 2008 Elsevier and 2012 American Chemical Society.

1.8.5 β-haematin Inhibition assays

The development of assays to measure β-haematin inhibition has been crucial for investigating the mechanism of action of chloroquine as well as other antimalarials. IR spectroscopy was the first method used in 1994 by Egan *et al* [79] in which the presence of free haem was reported using high concentrations of acetate as a methodology for formation of β-haematin crystals. Dorn *et al* [63], Sullivan *et al* [78] and Kurosawa *et al* [87] all used the same acetate methodology, albeit with lower concentrations of acetate and radio-labelled haematin. Slater *et al* [60] also showed that the acetonitrile extract of trophozoite lysate can induce β-haematin formation. However, these assays did not account for the role of lipids in

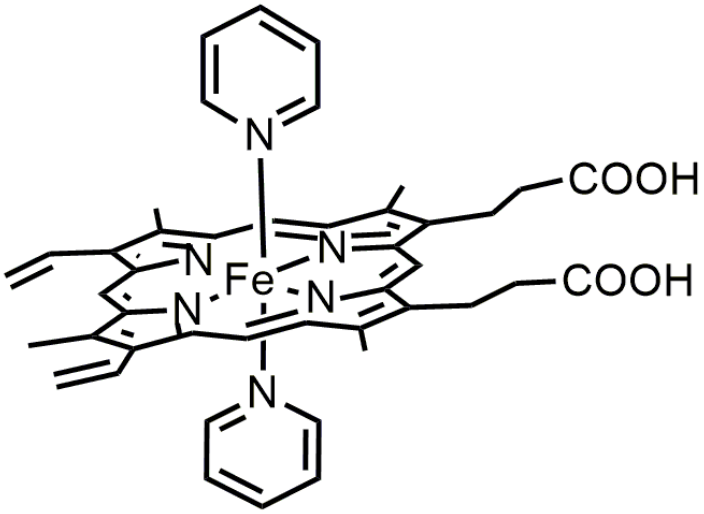
the formation of haemozoin and thus Fitch *et al* [88] reported the use of detergents and lipids to mediate the formation of β-haematin. The use of neutral lipid blends is time-intensive and costly thus in order to mimic lipids, the use of surfactants was attempted. The main characteristic of surfactants is that they contain a hydrophobic tail and a hydrophilic head that can be charged, uncharged or zwitterionic. They consists of a broad range of detergents, foaming agents, dispersants and emulsifiers. In water, these molecules aggregate forming micelles, a spherical shape where the hydrophobic tails are contained within the sphere away from the hydrophillic water and the head groups point towards the water. This is only possible if the surfactant has reached the critical micelle concentration otherwise the surfactant forms a layer above the water with the hydrophilic heads closest to the surface of the water.[89]

The initial attempt by Huy *et al* [90] in 2007 to use TWEEN® 20, a detergent, to mediate β-haematin formation had mixed results. The incubation time was found to be less and it was found to be cheaper than using neutral lipid blends. However, the results showed a ten-fold increase in the IC_{50} values of known antimalarials. Following this, Carter *et al* [91] reported the use of another detergent, Nonidet-P40 (NP-40) which had comparable results to that of the neutral lipid blend experiments. Along with NP-40, a range of detergents, namely, TWEEN® 20, TWEEN® 80, CHAPS, SDS and Triton X-100 were investigated. The detergents ability to nucleate crystal growth was investigated and the β-haematin products were characterised using PXRD and IR spectroscopy. The only β-haematin crystals that matched those of the Bohle dehydrohalogenation β-haematin synthesis (yields ≥ 69%) were formed using NP-40, TWEEN® 80 and TWEEN® 20. Further concentration dose response experiments using known antimalarial drugs demonstrated that only NP-40 showed comparable results to that of the neutral lipid blend methodology after the concentration of NP-40 was optimised. These results are shown in Table 1.1.

Table 1.1: Comparison of detergents using the data collected by Carter *et al* [91] where NP-40 shows the best yield and has a comparable IC_{50} value to that of the neutral lipid blend methodology.

Lipophilic Mediator	Mediator Concentration (μM)	β-haematin yield (%)	Chloroquine IC_{50} (μM)	Amodiaquine IC_{50} (μM)
Neutral Lipid Blend	50	75	23.07	85.26
NP-40	30.6	74	25.73	50.99
TWEEN® 80	14.9	69	201.8	195.7
TWEEN® 20	9.77	71	316.3	262
SDS	1140	10	225.4	245.6
CHAPS	1466	7.0	n/a	n/a
Triton X-100	58.6	7.4	n/a	n/a

This assay was originally designed to be used for a high-throughput screening study where the compounds in question contained a variety of functional groups and thus the assay needed to be robust. In 2011, this assay was adapted for use in 384-well plates by Sandlin *et al* [92] and the colorimetric pyridine ferrichrome method developed by Ncokazi and Egan [67] was used to detect the amount of haem present in each well. This method was originally adapted from a well known way of quantifying haem in haemoglobin in which pyridine is added to the mixture and forms a low-spin bispyridyl complex (Scheme 1.6). This assay also forms this complex with free haem but not β-haematin. The absorbance spectrum of the complex showed a shift of the Soret band from 389 to 404 nm and sharpened it while increasing the Q bands intensity which indicated a change in speciation from an aggregate (most likely dimers), which is known to occur with haem in water, to a low-spin monomer. It was also demonstrated that this obeyed Beer's law up to concentrations of 30 μM at 405 nm. The pyridine concentration and pH of the assay were optimised to allow only formation of the bispyridyl complex and allow no disruption of β-haematin which is seen at high pH.



Scheme 1.6: The bispyridyl-haem complex formed with the addition of pyridine to haem described by Ncokazi and Egan.[67]

1.9 High-throughput Screening

High-throughput screening (HTS) is the process whereby automation enables quick assay of biochemical or biological activity of a large number of drug-like compounds. It has been widely used in the pharmaceutical industry for the discovery of hit compounds from large libraries of compounds such as those of Novartis [93] and GlaxoSmithKline (GSK).[94]

1.9.1 Target-based Versus Phenotypic Screening

Three types of high-throughput screening are known, namely, virtual, phenotypic and target-based of which only target-based and phenotypic will be discussed for the purpose of this thesis.[95] Target-based screening is one in which, as the name suggests, the target of the compound is known and compounds are screened against that target. This has been the most popular form of screening especially for malaria based drug targets due partially to its reduced cost compared to phenotypic screening.[95,96]

Phenotypic screening aims to measure an effect in tissue or whole cells without any knowledge of targets. This type of screening suffers from high cost and additional work in terms of target identification. However, both these screening methodologies can be combined by firstly performing target-based screening on the whole library of compounds, followed by phenotypic screening to analyse additional factors that contribute to activity such as cell permeability.[95,96] Recently, there have been discussions to move back to phenotypic

screening due to improved target identification tools such as improvements in genetics and proteomics.[96]

1.9.2 Antimalarial High-throughput Screening

The initial HTS for β-haematin inhibitors was performed in 2000 by Kurosawa *et al*.[87] Several libraries of compounds were screened, containing over 100 000 compounds for their β-haematin inhibition activity which used the radio-labelled haematin methodology. From these 100 000 compounds only 45 non-quinoline compounds were identified with four of these compounds, a triarylcarbinol, a piperazine, a benzophenone and an unknown scaffold found to be active against the *P. falciparum* NF54 (chloroquine sensitive) and K1 (chloroquine resistant) strains.

Rush *et al* [97] also screened 16 000 compounds for β-haematin inhibition activity in 2009. However, the colorimetric pyridine ferrihemochrome methodology developed by Ncokazi and Egan [67] was used instead. This led to about 600 hit compounds of which 17 showed activity against both the 3D7 (chloroquine sensitive) and Dd2 (chloroquine resistant) strains. Both these screens described were target-based screens.

The largest antimalarial HTS by far was carried out by two industrial companies (Novartis and GSK) as well as an academic group (St. Jude Children's Research Hospital) using a phenotypic screening method on *P. falciparum* between 2008 and 2010. GSK screened about two million compounds from their libraries of which 13 533 where confirmed to inhibit parasite growth by more than 80% at a concentration of 2 μM.[94] This set of compounds became the Tres Cantos Antimalarial Set (TCAMS). The St. Jude Children's Research Hospital screened 309 474 compounds with the same cut-off of greater than 80% parasite growth inhibition against the 3D7 (chloroquine sensitive) and K1 (chloroquine resistant) strains showing 561 hit compounds.[98] Lastly, Novartis screened 1.7 million compounds of which 5 973 where found to active and novel enough to be considered as hit compounds.[93]

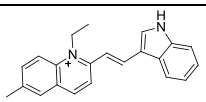
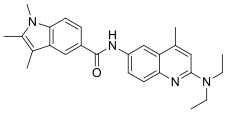
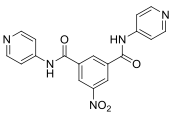
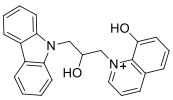
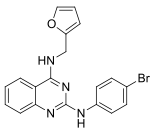
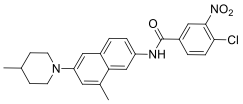
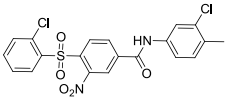
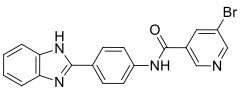
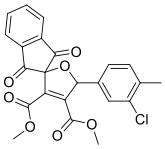
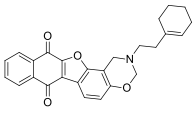
1.9.3 Vanderbilt University HTS Efforts

Following the initial pilot study containing 38 400 compounds from the Vanderbilt University library which was used to test the validity of the HTS methodology, a larger study was conducted using all of the compounds (144 330) to investigate novel β-haematin inhibiting

scaffolds. Compounds were first tested for β-haematin inhibition activity using the NP-40 detergent based assay in 384-well plates with positive control amodiaquine and negative control DMSO used. Compounds that were found to inhibit greater than 80% of β-haematin formation were considered hits which identified 729 compounds with a hit rate ([hit compounds/total compounds] x100) of 0.5%. Hit compounds were then subjected to a so called “cherry picking” methodology to test for false positives. After a final cut-off of 27 µM was established, 530 compounds were considered hits.

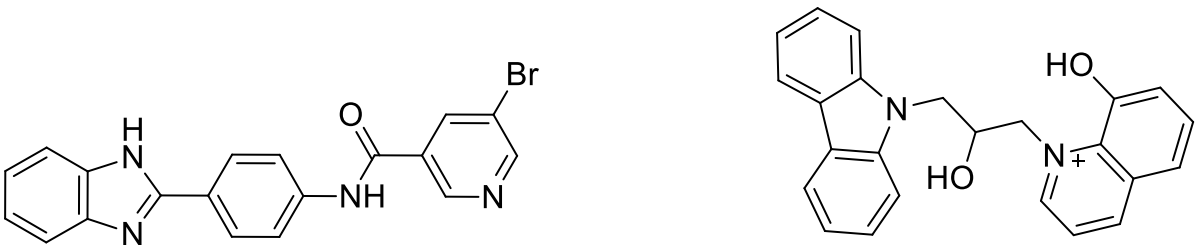
Following this, all hit compounds were tested for parasite growth inhibition activity against the D6 (chloroquine sensitive) strain using the SYBR Green I fluorescence-based assay first described by Smilkstein *et al*.[99] 171 of these hit compounds were shown to inhibit ≥ 90% parasitemia and considered hits with a hit rate of 32%. Dose-response analysis of these hits showed 73 compounds to have activities of ≤ 5 µM of which 25 exhibited nanomolar activity. These compounds that exhibited nanomolar activities were further tested on the C235 (chloroquine resistant) strain of the parasite using the same conditions as before. The resistance index of these compounds was calculated with 21 out of 25 showing values of less than 3 indicating effectiveness against resistant parasites.[100] The top ten of these compounds are shown in Table 1.2.

Table 1.2: The data collected by Sandlin *et al* [100] indicating the top 10 most active compounds against β-haematin formation and parasite growth inhibition in both the resistant and chloroquine sensitive strains found during the HTS.

VU Identifier	Structure	β-haematin inhibition IC_{50} (μM)	D6 IC_{50} (μM)	C235 IC_{50} (μM)	Resistance Index
VU0098755		12.6	0.11	0.13	1.2
VU0073687		6.3	0.18	0.55	2.9
VU0001281		5.7	0.19	0.17	0.9
VU0065708		16.2	0.20	0.18	0.9
VU0096505		8.8	0.24	0.22	0.9
VU0107282		17.0	0.29	0.54	1.9
VU0114785		13.4	0.35	4.82	13.9
VU0002101		14.3	0.35	0.41	1.2
VU0028177		13.3	0.35	0.46	1.3
VU0063971		8.9	0.38	0.83	2.2

Further analysis of a respresentative of each scaffold found in the HTS was subjected to the cellular haem fractionation assay to determine if haemozoin inhibition is the correct mode of action for these compounds. This assay was developed by Combrink *et al* [71] and measures the percentage free haem and percentage haemozoin at various concentrations of drug. A haemozoin inhibitor will show an increase in free haem and decrease in haemozoin. All of the hits tested exhibited positive results indicating that all the hits were haemozoin inhibitors.[100]

Fourteen different β-haematin hit scaffolds were found in this study, 46% of which encompass the total β-haematin inhibitors in the study. Of the fourteen scaffolds two were analysed more closely, namely, the benzimidazoles and the carbazole scaffolds. The HTS results showed ten benzimidazole compounds to be active against β-haematin formation of which five where found to be active against the parasite. Similarly, the carbazole scaffold showed eight β-haematin inhibitors, four of which where found to be active against the malaria parasite. These scaffolds are also represented in the top 10 most potent hits shown in Table 1.2 with ranks 4 and 8 representing carbazoles and benzimidazoles respectively.[100] Therefore those hit compounds were chosen for analysis in this thesis and shown in Scheme 1.7.

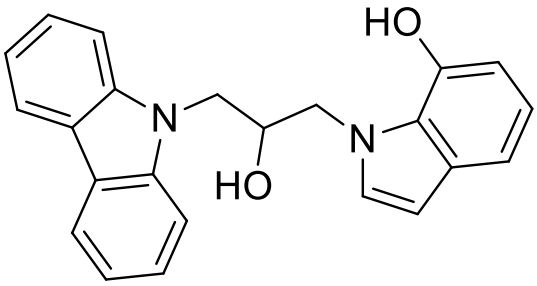


Benzimidazole Hit **Carbazole Hit**

Scheme 1.7: The two scaffolds chosen from the HTS study for further analysis.

Due to the unfavourable positive charge on the carbazole hit, it was decided to modify this hit using bioisosteres. Using the online resource, SwissBioisoteres, enabled the choice of indoles instead of the quinoline group. Swissbioisosteres indicated that 17 out of the 19 compounds on record with this ring system change retained the same activities, while two showed a better activity and one showed a lower activity. Therefore, the quinoline group was

changed to an indole group as shown in Scheme 1.8. The following sections will discuss these scaffolds in reference to known literature drugs.



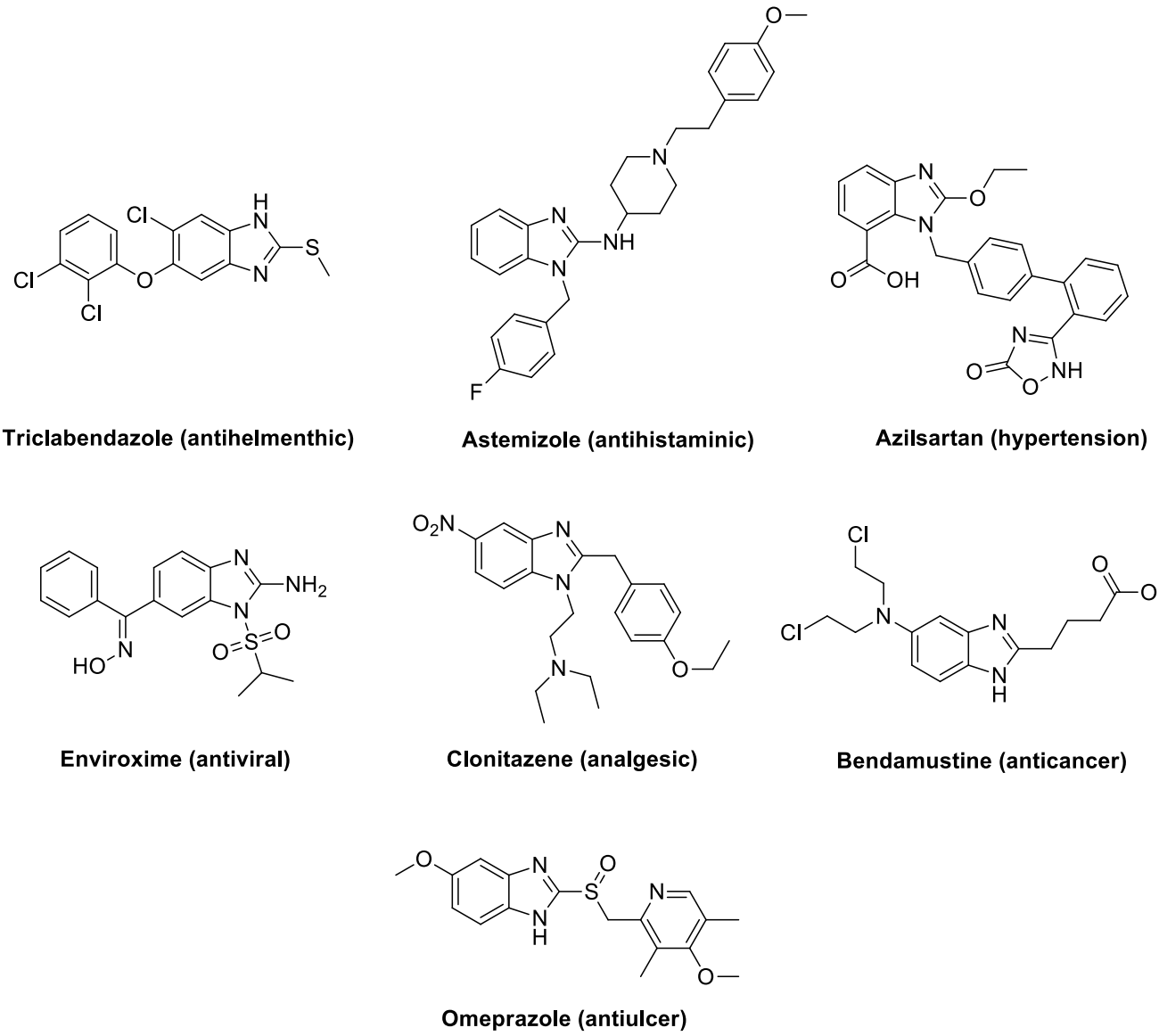
Scheme 1.8: New proposed carbazole based hit analogue.

1.10 Benzimidazole Scaffold in the Literature

Benzimidazoles are a well-known chemical moiety that was first synthesized by Hoebrecker and subsequently by Wundt and Landenberg from the years of 1872 to 1873 which now forms part of a diverse family of bioactive compounds that have a broad range of medicinal uses.[101–103] These medicinal properties include, but are not limited to, antiparasitic, antimicrobial, potential anticancer, anti-inflammatory, analgesic, antihypertensive, antioxidant, antiulcer, anti-HIV and antimalarial activity amongst others (Scheme 1.9).[103–111] This wide variety of medicinal applications explains the prevalence of interest in this scaffold as well as syntheses involving the benzimidazole scaffold throughout literature.[103]

Since the late 1980s, there have been numerous reports of benzimidazole compounds inhibiting the growth of *Plasmodium* parasites. In 1990, Dieckmann-Schuppert and Franklin [112] reported benzimidazole compounds that interact with tubulin to inhibit tubulin polymerisation. The same authors also reported attempts at using albendazole, a known benzimidazole drug that is used in the treatment of hookworms, against the malaria parasite. This was not feasible as high concentrations of the drug were needed to be effective.[113] Mayence *et al,* [113] in 2011 reported bis(oxyphenylene)benzimidazole compounds that were found to be active against *P. falciparum.* Similarly, there were reports of amidino bis-benzimidazoles and *N*-aryl-2-aminobenzimidazoles reported by Worachartcheewan *et al* [114] and Ramachandran *et al* [115] respectively that demonstrated activity against *P. falciparum*. However, none of these reports show haemozoin inhibiting benzimidazole compounds apart

from the recent work by Chibale *et al* [116] which showed weak β-haematin inhibitory activities. This leaves an interesting research gap for the benzimidazole hit compound shown in Scheme 1.7.

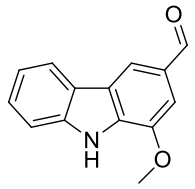


Scheme 1.9: Some prominent drug compounds containing the benzimidazole scaffold indicating their wide medicinal applications.[117]

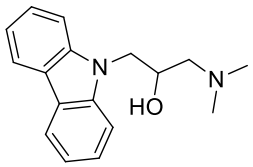
1.11 Carbazole Scaffold in the Literature

Carbazole was first isolated by Graebe and Glazer in 1872 from coal tar.[118] In 1965, the first carbazole alkaloid, murrayanine, was isolated from the leaves of the *Murraya koenigii* Spreng tree in India.[119] Following this, there has been a keen interest in carbazole-containing

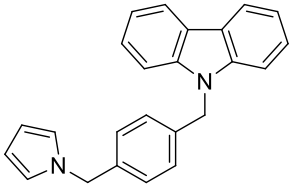
compounds due to their various biological applications which mainly include anticancer [120,121], antimicrobial [122,123] , neuroprotective and recently antimalarial activity.[124] Scheme 1.10 shows some compounds mentioned in literature that have potent activities against their relevant targets.



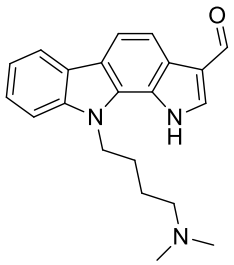
Murrayanine



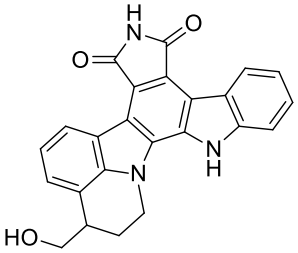
Wiskostatin (antimicrobial)



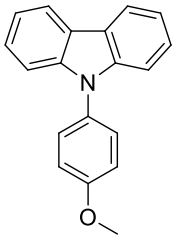
Carbazole imidazole (antimicrobial)



***N*-substituted pyrrolocarbazole (anticancer)**



***N*-annulated indolocarbazole (anticancer)**

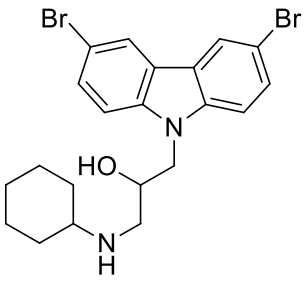


Phenyl carbazole (neuroprotective)

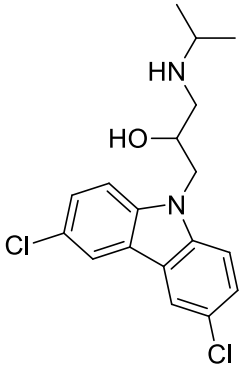
Scheme 1.10: Carbazole-containing compounds that showed potent activities for treatment of various diseases.

Shortly after murrayanine was discovered and synthesized, investigations into the carbazole-containing compounds for antimalarial studies were carried out. Brown *et al* [125] investigated the potential for heteroaryl carbinols to exhibit antimalarial activity in 1971, but all the compounds were shown to be inactive. Following this in 2007, a patent was filed for the discovery of novel carbazole-containing compounds that were shown to inhibit haemozoin formation via the supposed inhibition of the haem detoxification protein which is theorized to catalyse haemozoin formation in the parasite. Molette *et al* [122] demonstrated the activity of a series of amino-alcohol carbazoles, similar to the hit compound found during HTS, with promising *in vitro* and *in vivo* results. The most active of those compounds is shown in Scheme 1.11. Wang *et al* [126] performed a HTS on a library of 126 compounds docking them *in silico* in

the *P. falciparum* specific binding site, *Pf*Hsp90. The best compound found during this screen was an amino-alcohol carbazole compound shown in Scheme 1.11. This compound showed good binding to the amino acid residues while showing no specific binding to the human analogue of *Pf*Hsp90. This compound also showed potent activity against *P. falciparum*.



Amino-alcohol carbazole (Molette *et al)*



Amino-alcohol carbazole (Wang *et al)*

Scheme 1.11: The amino-alcohol carbazole derivatives that were demonstrated to be the most active against *P. falciparum* in each literature study.

Given that the benzimidazole and carbazole hit compounds chosen from the HTS are backed up by a rich source of data from literature, indicating their potential as druggable compounds together with their potent activities shown in the HTS, these compounds were considered ideal candidates for further investigation in this project.

1.12 Aims and Objectives

1.12.1 Aims

Due to the dire need for new scaffolds of antimalarials that are able to overcome growing resistance, while retaining haemozoin formation as the drug target of choice, the aims of the project were to synthesize derivatives of the hit compounds chosen from the Vanderbilt HTS using Bayesian statistical models to narrow down the choice of compounds, probe the mechanistic action of these derivatives and discover any possible structure-activity relationships.

1.12.2 Objectives

More in depth lists of objectives for each chapter are shown in the introduction of each chapter with a summarised version shown below.

In order to achieve the aims of this project the following specific objectives needed to be achieved:

- Enumerate derivatives of the benzimidazole scaffold and predict β-haematin inhibition and malaria parasite growth inhibition values using Bayesian statistics for both scaffolds.
- Select derivatives for synthesis based on these results and synthesize a series of derivatives for each scaffold.
- Analyse all the synthesized compounds for biological activities and compare it to the predicted Bayesian results.
- Perform molecular docking on all compounds to determine any correlations with β-haematin inhibition activity.
- Perform the cellular haem fractionation to verify the haemozoin target of these scaffolds.
- Investigate all SAR of these compounds to enable further optimisation of these scaffolds in the future.

Chapter 2: Bayesian Studies

2.1 Introduction

HTS has been one of the primary tools used by the pharmaceutical industry for years as it uses large libraries screened against single targets to discover hits. Despite its success in the industry there is always a search for a more cost effective way to determine the appropriate chemical characteristics, pharmacokinetics and pharmacodynamic properties of potential hits.[127] HTS also suffers from lack of chemical diversity as large databases generally only focus on a small chemical space.[128] Machine learning techniques, such as Bayesian classifiers and molecular docking helps to screen a whole array of compounds *in silico* and predict their activities as well as give a better understanding of the interaction of these compounds with targets such as the β-haematin surface. This allows focus to be drawn onto different chemical spaces, eliminating most of the false-positives that are found in HTS screens and shortens time needed for synthesis.[129,130]

2.2 Bayesian Probability Method

2.2.1 Bayesian Statistics

Machine learning approaches have been used for a variety of chemoinformatic applications in the literature, including the aiding of drug design by predicting molecular properties or compound activity.[131,132] Some of the more frequently used statistical algorithms include random forests, vector machines and naïve Bayesian, however, naïve Bayesian has only recently been applied to antimalarial activity predictions.[127,131] Bayesian statistics can be used to build a statistical model from a large data set of compounds with both active and inactive compounds present such as from a HTS data set (priors) to predict properties such as adsorption, distribution, metabolism and excretion (ADME) as well as activities against the malaria parasite.[127,132,133]

2.2.2 Bayes' Theorem

The statistical model used in this study makes use of Bayes' theorem, first published in 1763 by Thomas Bayes, to predict molecular properties based on Equation 2.1: [134]

$$P(A|M) = \frac{P(M|A)P(A)}{P(M)}$$ **Equation 2.1**

For the purpose of this thesis, A is the activity of a compound and M is the molecular feature that contributes to the activity. Therefore P(A) is the probability of any compound being active in the training set, P(M) is the probability of the molecular feature being present in any molecule in the training set, P(A|M) is the probability that the molecular feature contributes to activity in a test molecule and P(M|A) is probability that the molecular feature is present in an active compound in the training set.

2.2.3 Bayesian Model in Discovery Studio

The Bayesian model approach in Accelrys Discovery studio software uses prior data, in this case HTS data, with knowledge of inactive and active compounds according to a predetermined cut-off set by the user to create a training set.[127,135,136] Any compound that is below the user determined threshold is earmarked as "bad" (inactive) and any compound above the threshold is earmarked as "good" (active).[127,129,133,137] The system learns to distinguish features of active compounds using input descriptors such as LogP, molecular weight, number of hydrogen bond donors, number of hydrogen bond acceptors, number of rotatable bonds, number of ring systems, fractional polar surface area, extended connectivity fingerprints of depth 6 and the number of aromatic rings.[127,131,137] The model records the frequency of these features in the '"good" compounds and assigns them a weighting using the Laplacian-adjusted probability estimate (see Section 2.2.4).[127,133,137] A Receiver Operating Characteristic (ROC) score is given to the user once the model is completed which indicates the performance of the model and is calculated from the area under the curve of the true positives rate versus the false positives rate. A ROC score of 1 means a perfect model and that of 0.5 means that no enrichment is achieved.[127,129,137] An external compound (test compound) can be run in the model and its data can then be added to the model to improve the ROC score. This can be done several times with active and inactive compounds as the model improves with a more diverse array of compounds.[137] For this work a prebuilt β-haematin inhibition model (cut-off 100 µM) and a prebuilt parasite growth inhibition model (cut-off 2 µM) with ROC scores of 0.915 and 0.991 respectively were used which was reported by Wicht *et al.*[127]

2.2.4 Laplacian-corrected Probability Estimator

In order to adjust for an uncorrected probability of a certain feature of a molecule, a Laplacian-corrected estimator is used as some molecular features appear more often than others which can skew the data. The derivation is given by means of an example from Xia Xiaoyang *et al* and Discovery Studio.[135,137]

Assuming a number of compounds T is available in a given training set of which X are active then the baseline probability of a random compound chosen being active would be $P(active) = X/T$. If Y of those T compounds contained a certain molecular feature Z and only S of those molecules were shown to be active, then the probability of a compound being active changes to $P(active|\mathrm{Z}) = S/Y$. The estimate for that feature Z becomes unreliable if the number of molecules in the training set with the feature Z is low. This means, for example, if only one compound in the training set contains molecular feature Z and happens to be active then $P(active|\mathrm{Z}) = \frac{1}{1} = 1$, which is vastly overconfident for that molecular feature. This means that the feature has been undersampled, but most molecular features do not relate to active molecules in the training set so for these cases $P(active) = P(active|\mathrm{Z})$.

Adding a sample of N molecules with the molecular feature Z to the training set, we now have the expected number of molecules being active is $P(active) * N$. With the inclusion of these new molecules to the training set, correcting the estimate of the probability, the equation changes to:

$$Pcorr(active|\mathrm{Z}) = \frac{(S+\mathrm{P(active)}*\mathrm{N})}{Y+N}$$

This correction factor now stabilises the estimator as the number of compounds Y+N approaches N, the probability of the molecular feature now converges to just P(active) which is the expected value of most of the features as most features do not contribute to activity in a molecule. Finally, the estimator is divided by P(active) and then the logarithm is taken thereafter to produce Equation 2.2.

$$\log Pfinal(active|\mathrm{Z}) = \log \frac{Pcorr(active|\mathrm{Z})}{P(active)} \qquad \textbf{Equation 2.2}$$

This is done to normalise the equation and to completely predict the probability of the molecule being active from the molecular features. Most molecular features will give a

$LogP_{final}$ that is approximately zero but some features that are common in active molecules $LogP_{final}$ will be positive. Similarly, if the molecular features are common in inactive compounds in the training set $LogP_{final}$ will be negative.[135,137]

2.2.5 Pipeline Pilot

Pipeline Pilot was developed in 2004 by SciTegic, later a subsidiary of Accelrys and was the first software that allows its user to fully automate analysis, data management and reporting protocols at the click of a button. This is done by a number protocols linked together as the user sees fit, to create a workflow of jobs which can be reused by other users later or extended to include more processes.[138] Pipeline Pilot is compatible with all functions available in the Discovery Studio package and was used for this study.

2.3 Bayesian Goals

The goals of this chapter were to:

- Produce a workflow of protocols in Pipeline Pilot to:
  - Enumerate all possible *in silico* benzimidazole compounds from the available starting materials in the Sigma-Aldrich catalogue.
  - Predict β-haematin inhibition and hence select active compounds *in silico*.
  - Predict the resulting active compounds using the malaria parasite growth inhibition model and select the active compounds *in silico*.
  - Select the active compounds according to predicted aqueous solubility.

- Investigate six proposed carbazole indole derivatives *in silico* using both the β-haematin inhibition model and malaria parasite growth inhibition model.

2.4 Results and Discussion

2.4.1 Benzimidazole Bayesian Methodology

The benzimidazole hit compound from the HTS has three possible points of derivatization (Figure 2.1) that were explored in this study. The three ring systems in Figure 2.1 (red, green and blue) can lead to an enormous number of enumerated *in silico* compounds, thus it was decided to use the Sigma-Aldrich database of available chemicals as a basis to explore *in silico* derivatives.

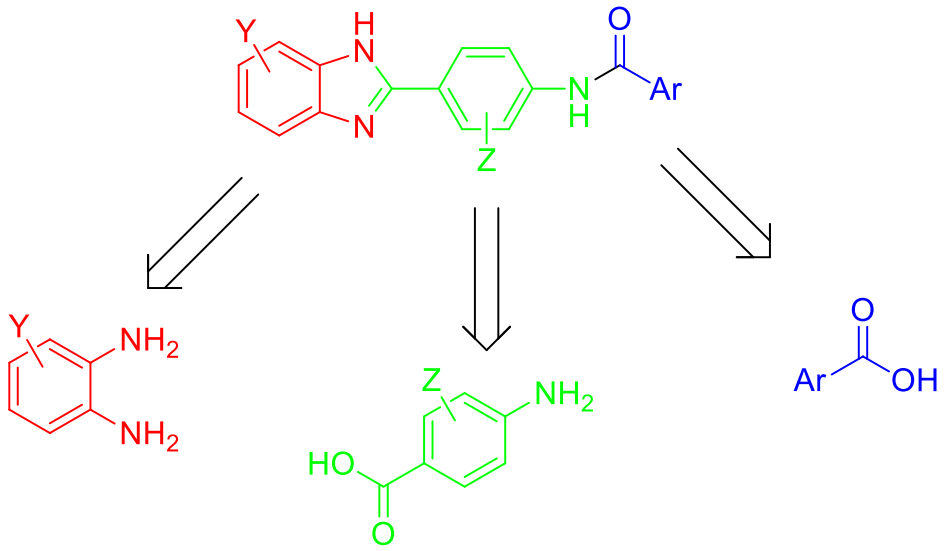
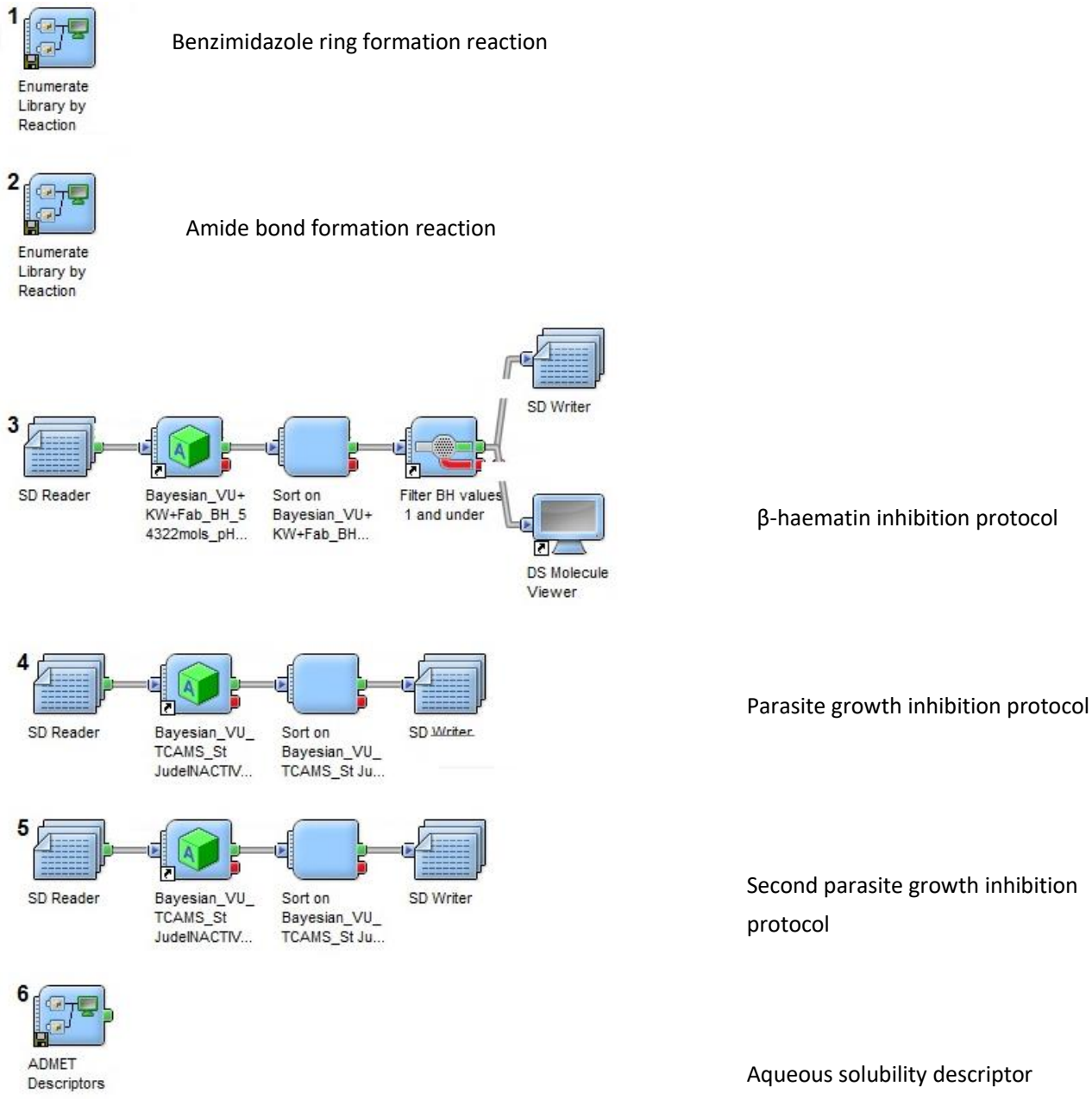


Figure 2.1: The three proposed areas of derivatization.

A protocol created in Pipeline Pilot (Scheme 2.1) was used in Discovery Studio to automate all the processes described below. This made it possible to create a workflow to narrow down the search for the optimum series of compounds to synthesize by using various filters after each step. Using this methodology 325 728 compounds were enumerated through a 2-step reaction from an input of 26 *o*-phenylenediamine derivatives, 24 *p*-aminobenzoic acid derivatives and 522 aromatic carboxylic acids. After enumerating this new *in silico* library of compounds, the next step in the workflow was to subject them to the β-haematin inhibition model created by Wicht *et al*.[127] All predicted inactive compounds (19 904), determined by the model cut-off of -4.920, were filtered out and the predicted active compounds (305 824) were filtered using the parasite growth inhibition activity model with a cut-off of -8.580 also

created by Wicht *et al*.[127] These predicted inactive compounds (2 148) were filtered out and the predicted active compounds (303 676) were sorted by aqueous solubility level with a cut-off of 1 or less, available as part of the Discovery Studio ADMET predictions function. 35 124 Compounds were predicted to have a moderate to low solubility whilst 268 552 were predicted to have little to no solubility in aqueous medium at 25 °C. Only 68 compounds of the 35 124 compounds were predicted to have a solubility value of 3 (moderate). This was a warning of the limitations of these compounds with any aqueous based chemistry and any assays that might need to be done on this series.



Scheme 2.1: An example of the Pipeline Pilot protocol used.

2.4.2 Benzimidazole Derivatives: Rationalisation

Reviewing the soluble active dataset, absence of substitution on the middle ring system (green in Figure 2.1) was found to be a defining feature. The only deviation from this result was the presence of mono nitro substituted derivatives and a change from a benzene to pyridine ring system. Upon analysis of the predicted β-haematin inactive compounds, 100% of the compounds in that subset of data contained one to four substituents on that ring system. In light of these results it was decided to avoid substituents on that ring system and focus on the other two positions of the molecule. This narrowed down the search even further and ensured a simpler synthesis of these compounds.

The other dominant features found in the soluble active compounds subset was the presence of mono and disubstituted nitro and halogen benzimidazole substituents, the presence of 5-membered ring systems after the amide bond (blue in Figure 2.1) and mono and disubstituted benzene and pyridine ring systems (blue in Figure 2.1). Selection of compounds with these features allowed for concentration of effort on the effect of the position of substitution on the various ring systems. It is also common practise to synthesize the hit compounds from the HTS study to ensure that this hit was not a false positive as all the data from the HTS was input into the Bayesian models so any compounds with a similar fingerprint would be heavily influenced by the hit and termed as active.

From these observations, the number of choices of compounds to synthesize was narrowed down from 35 124 to 36. Due to unforeseen synthetic difficulties described in Chapter 3, formation of compounds substituted on the benzimidazole ring (red in Figure 2.1) was not possible, therefore only the amide linked ring system (blue in Figure 2.1) was derivatized. Table 2.1 contains all the compounds chosen for this study and their appropriate predicted β-haematin and parasite growth inhibition activity. These compounds were predicted to have a lower aqueous solubility than their nitro and halogenated analogues, so it was expected that solubility issues would arise with these compounds, but their β-haematin inhibition and parasite growth inhibition scores far surpassed the cut-off values for these models and so they were considered ideal candidates for this study.

Table 2.1: Selected compounds for synthesis based on Bayesian results

Ar (compound number)	β-haematin inhibition score	Parasite growth inhibition score
Phenyl (**2**)	25.6979	61.3259
m-Bromophenyl (**3**)	28.1166	57.2935
m-Trifluoromethylphenyl (**4**)	32.6362	75.2685
3-Bromopyridyl (**5**)	38.4248	72.6119
p-tert Butylphenyl (**6**)	25.4777	63.8076
2,4-Dichlorophenyl (**7**)	24.0156	49.3979
o-Methoxyphenyl (**8**)	19.9991	43.3265
3-Trifluoromethylpyridyl (**9**)	36.4715	81.8491
o-Nitrophenyl (**10**)	21.0876	30.0400
3,5-Dinitrophenyl (**11**)	40.5853	47.2409
2,6-Dichloropyridyl (**12**)	29.5242	65.4908
o-Chlorophenyl (**13**)	19.5834	45.2802
m-Chlorophenyl (**14**)	31.0259	58.6210
p-Chlorophenyl (**15**)	25.1454	56.5011
3-Pyridyl (**16**)	30.6674	78.3569
Thiophenyl (**17**)	19.3296	55.1323
Furanyl (**18**)	14.9506	48.2215
Pyrrole (**19**)	27.4973	67.9852

2.4.3 Benzimidazole Fragments

An *in silico* investigation of the hit benzimidazole fragments was performed using the previously described Bayesian models to predict their activities. This provided a rationalization for the change in predicted activity with the removal of key functional groups. The compounds derived from the hit compound are shown in Figure 2.2 and are comprised of the smallest fragment on either side of the molecule built sequentially until the compounds lack just one ring system. The Bayesian statistics results for the β-haematin model showed all but one compound to be predicted as active but many compounds were found to be close to the cut-off value as shown in Table 2.2. This prediction indicated that these compounds would either be expected to have low, or no β-haematin inhibition activity.

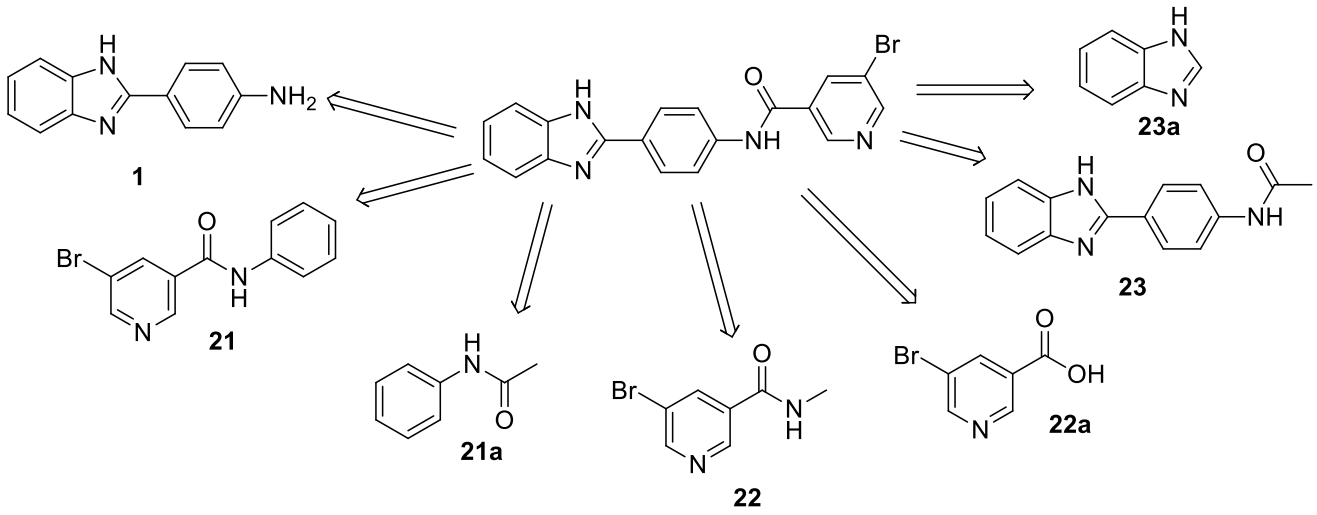


Figure 2.2: Benzimidazole fragment compounds derived from the hit compound in the centre.

The results of the parasite growth inhibition model predicted that five out of the seven compounds would be active, with three of the compounds close to the cut-off indicating high confidence in just two predicted active compounds. Both compounds **1** and **23** contained the benzimidazole and middle benzene ring system, supporting the previous decision to focus synthetic efforts on the carbonyl containing ring system.

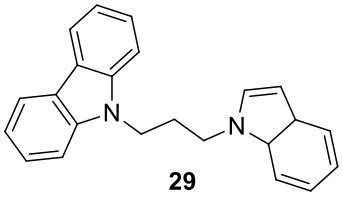
Table 2.2: The Bayesian statistic results of the benzimidazole fragment study with predicted active compounds shown in red and predicted inactive compounds shown in black.

Compound Number	β-haematin inhibition score	Parasite growth inhibition score
1	22.8883	56.9520
21	12.7431	-8.57182
21a	-5.34745	-10.6894
22	6.81007	-8.25818
22a	8.04823	-12.8326
23	19.8929	60.8754
23a	1.03421	4.55447

2.4.5 Carbazole Indole Derivatives

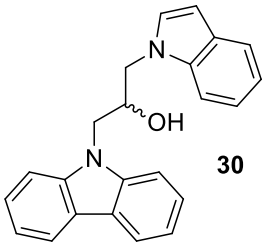
Bayesian statistics can provide some insight into the role of hydrogen bonding of the hit compound with the β-haematin crystal surface. Enumerating compounds by sequentially removing hydroxyl groups from the hit compound until it was devoid of hydroxyl groups is a case in point with respect to the carbazoles (Figure 2.3). This generated only three different compounds, but using the hit data it was unclear which enantiomer was active therefore all enantiomers and racemic compounds were enumerated for this study.

Compound with no hydroxyl groups

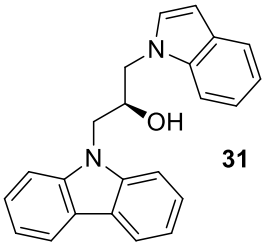


29

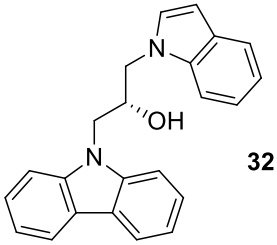
Compounds with one hydroxyl group



30

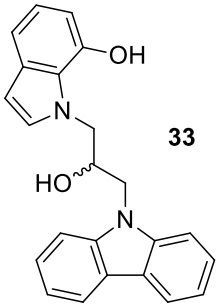


31

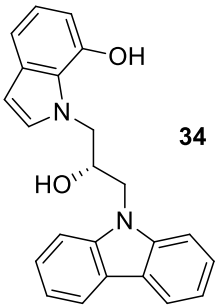


32

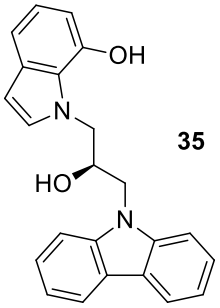
Compounds with two hydroxyl groups



33



34



35

Figure 2.3: Enumerated carbazole indole compounds from the neutral hit analogue (**33)**.

Predicted activities (Table 2.3) were computed using both the Bayesian β-haematin inhibition and the parasite growth inhibition model, as with the benzimidazole compounds. Interestingly, only compounds **33**-**35** were predicted to be active for the β-haematin inhibition and compounds **30**-**35** where predicted to be active in the parasite growth inhibition model however, the numerical values are low for these compounds termed to be active. Remembering that the cut-off of the β-haematin inhibition model is -4.920, the difference of the active values is 9.645 units and with the parasite growth inhibition model cut-off value of -8.580, the difference is 3.198 and 0.386 units. This implies that the data is unreliable and thus warrants further investigation.

Table 2.3: The Bayesian statistic results of the carbazole indole study using both models with predicted active compounds shown in red and predicted inactive compounds shown in black.

Compound Number	β-haematin inhibition score	Parasite growth inhibition score
29	-7.4793	-10.3333
30	-5.29619	-8.19434
31	-5.29619	-8.19434
32	-5.29619	-8.19434
33	4.72502	-5.38157
34	4.72502	-5.38157
35	4.72502	-5.38157

One drawback found using Bayesian predictions for this subset of compounds is that the models cannot distinguish any differences between these chiral molecules and thus the values for the racemic mixture and both enantiomers are identical. This is because the Bayesian models use sets of molecular fingerprints to predict activity, but in the Bayesian model used there were no molecules with a chiral fingerprint so it cannot differentiate between these chiral centres.

2.5 Summary and Conclusions

In this study 325 728 unique benzimidazole compounds were enumerated using the Sigma-Aldrich catalogue of starting materials with the hit benzimidazole compound a basis for derivatization. These compounds underwent two stages of activity predictions using pre-existing Bayesian models, filtering out predicted inactive compounds at each stage. 19 904 compounds were predicted to be inactive in the β-haematin inhibition model from the initial 325 728 compounds, after which 2 148 compounds were predicted to be inactive in the parasite growth inhibition model. The 303 676 predicted active compounds underwent an aqueous solubility prediction producing 268 552 moderately soluble compounds. 18 of these compounds were selected for synthesis based on these results and it was decided to

derivatize only one ring system. The hit benzimidazole subunits were also investigated in a similar fashion showing six out of seven compounds predicted to have β-haematin inhibitory activity while five out of the seven compounds were predicted to have parasite growth inhibitory activity. This was also carried out on seven carbazole indole compounds, predicting four out of seven compounds to be inactive against the β-haematin inhibition model and one out of seven compounds inactive against the parasite growth inhibition model.

Implementing the Bayesian statistical techniques in this study has greatly enriched the knowledge base on this scaffold regarding choices and positions of functional groups. Bayesian statistics has also improved our throughput, enumerating 325 728 compounds *in silico* without the need for synthesis. As with all techniques, there are limitations with this methodology as the results depend on the training set of each model and its ability to recognize certain fingerprints, thus from this study the enantiomers of the chiral carbazole indole scaffold were predicted to be the same. Further investigation into these scaffolds was considered necessary using experimental and other computational techniques.

Chapter 3: Synthesis Aspects

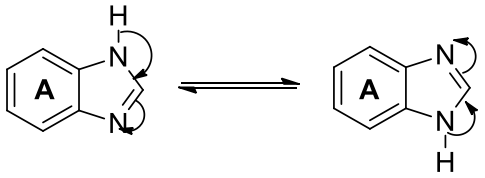
3.1 Introduction

The Bayesian predictions in Chapter 2 invited the synthesis of the compounds identified, in which first a retrosynthetic analysis of both the hit compounds as well as the relevant literature methodologies for the transformations involved and their mechanisms will be discussed. Thereafter, the actual synthesis results with the appropriate characterisation techniques will be discussed, which will be followed by an extension into series generation for structure-activity studies. The derivatives synthesized were not necessarily meant to be drug candidates but ones to reinforce knowledge of the mechanism of action of these compounds and to validate previously discussed predicted activities from the Bayesian statistics.

3.2 Scaffold 1: Benzimidazoles

3.2.1 Benzimidazole Chemical Properties

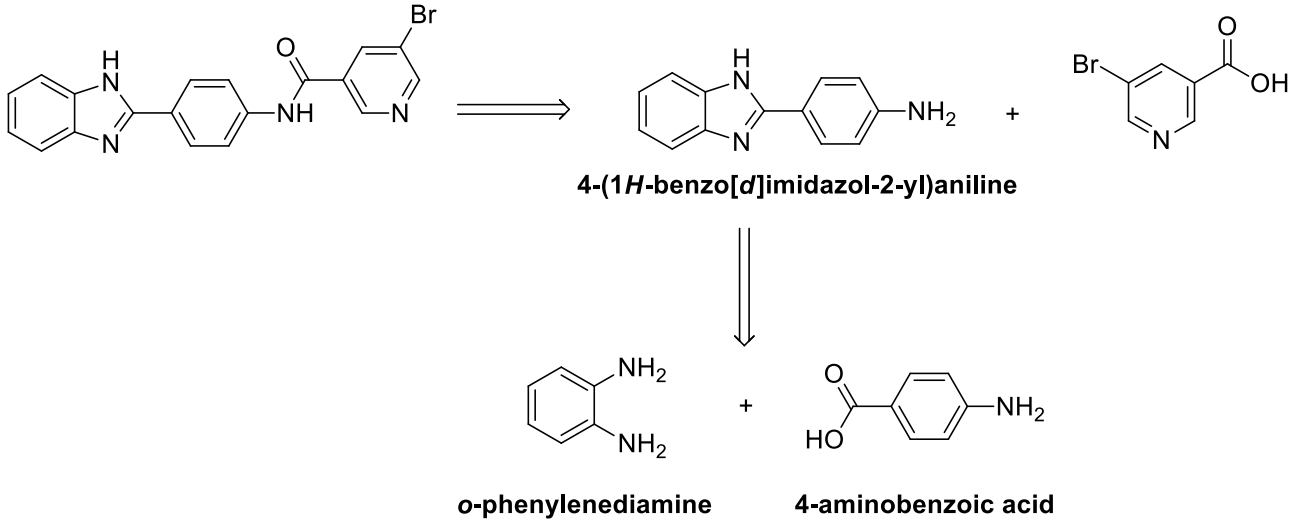
The benzimidazole moiety has two nitrogen atoms contained as an activated imine and a mildly acidic amine due to resonance connection between the two. Both functionalities are involved in H-tautomerism (Scheme 3.1) and in water this process is reported[139,140] as being an intermolecular process between the benzimidazole imine of one molecule and the benzimidazole amine of another molecule, which results in the nitrogen atoms becoming chemically equivalent. Any mono-substitution on the A-ring, however, introduces an asymmetry that results in a non-equivalence in the H-tautomers. Similarly, *N*-substitution with A-ring mono-substitution also results in constitutional isomers being produced.[140]



Scheme 3.1: Tautomerism of the benzimidazole ring system.[140]

3.2.2 Benzimidazole Hit Compound Retrosynthetic Analysis

The benzimidazole hit compound can be disconnected back to a 4-(1*H*-benzo[*d*]imidazol-2-yl)aniline intermediate which can be further disconnected back to *p*-aminobenzoic acid and *o*-phenylenediamine as shown in Scheme 3.2. The forward reactions therefore consisted of two separate condensation reactions to form the benzimidazole ring and the amide bond respectively. These observations led to a literature survey of both of these reactions, which are now reviewed in context.

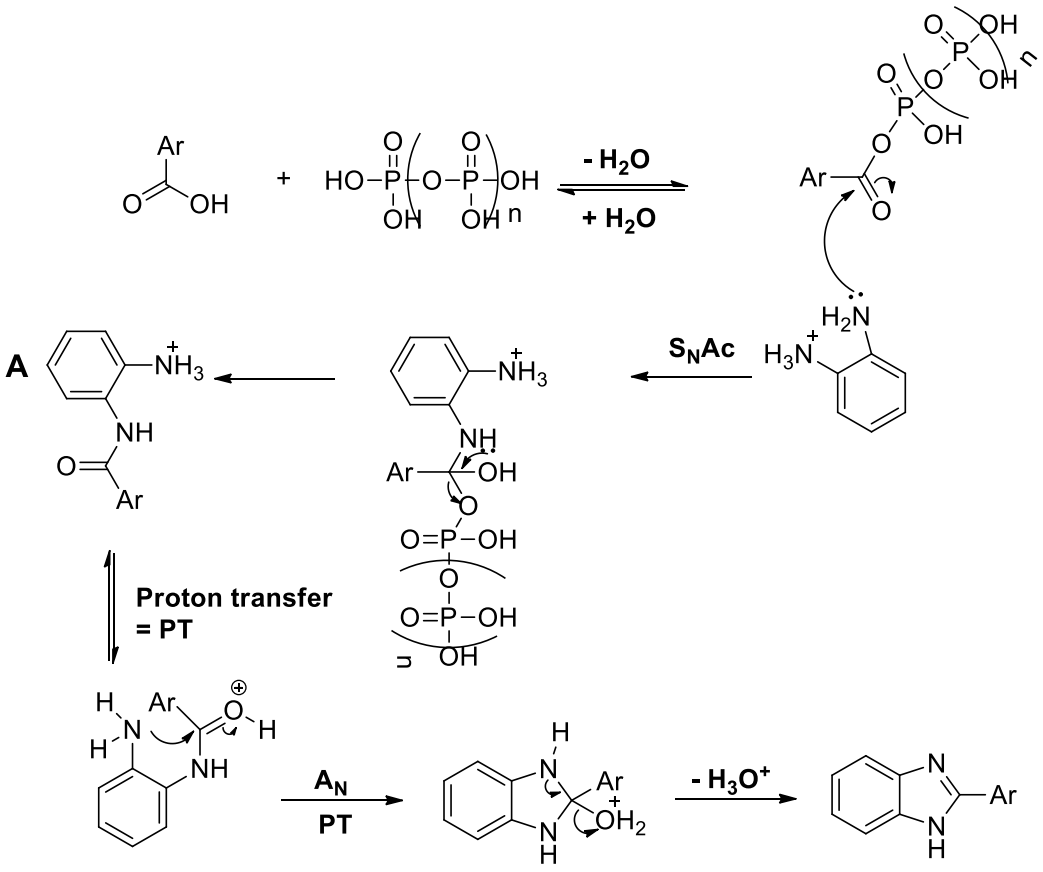


Scheme 3.2: Retrosynthetic analysis of the benzimidazole hit compound.

3.2.3 Benzimidazole Ring Synthesis

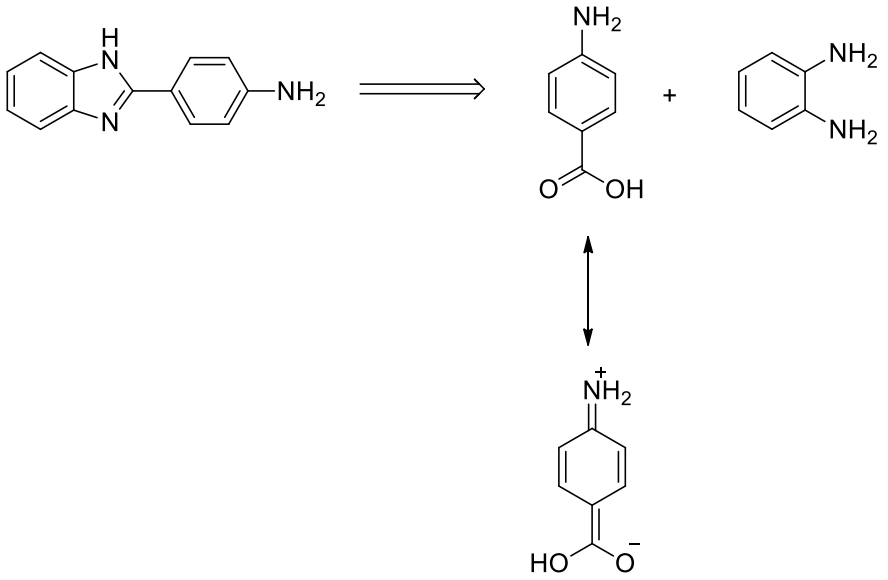
The benzimidazole ring system is traditionally prepared using the Phillips-Ladenburg reaction, which involves reacting a 1,2-diaminobenzene with a carboxylic acid using aqueous hydrochloric acid at high temperature or pressure.[141,142] This method gives poor yields when preparing 2-arylbenzimidazoles from aromatic carboxylic acids, which led Leavitt, Hein and Alheim in 1956 to improve the reaction using polyphosphoric acid (PPA) as an alternative promoter. This in turn resulted in a significant increase in yield, making it still one of the main methods used today for this target.[143,144] Modern approaches have expanded on this to include various Lewis acids, but in the present work it was decided to use the methodology laid out by Leavitt *et al* in view of its simplicity for accessing bulk quantities of the desired benzimidazole, together with the fact that the reagents were already available in-house.[103,144]

So and Heeschen[145] have shown by various nuclear magnetic resonance (NMR) studies that the mechanism for the formation of the benzimidazole ring system firstly, involves the formation of a mixed anhydride from the aromatic acid and PPA, which forms a dynamic equilibrium with the free benzoic acid group and PPA polymer such that the proportion of mixed anhydride increases with temperature. This can be visualised via NMR spectroscopy. It should be also noted that the percentage of mixed anhydride increases with initial phosphorus pentoxide (P_2O_5) content in the PPA, in which a P_2O_5 content of 77% or greater was needed for any mixed anhydride to be formed. Under these acidic conditions, the *o*-phenylenediamine exists as an equilibrium between its protonated and neutral forms (Scheme 3.3), the latter being the reactive species that combines with the mixed anhydride in a nucleophilic acyl substitution reaction (S_NAc) to form intermediate **A**.[145] The acidic medium promotes ring closure via another S_NAc mechanism involving nucleophilic attack of the other primary amine, which is also unprotonated, on the electrophilic carbon of the protonated carbonyl group. Finally, elimination of water furnishes the 2-arylbenzimidazole.[145] Scheme 3.3 below summarises these events:



Scheme 3.3: Reaction mechanism for formation of the benzimidazole product.

The key intermediate for producing the first series of compounds was the 2-arylbenzimidazole intermediate containing a *para* amino group in the 2-phenyl group, which could be disconnected back to *p*-aminobenzoic acid and *o*-phenylenediamine as shown in Scheme 3.4. The PPA methodology described above was used to form this intermediate. The amino group of the *p*-aminobenzoic was considered to be sufficiently deactivated nucleophilically due to resonance stabilisation by the carboxylic acid as shown in Scheme 3.4 and hence wasn't protected first. This would have lengthened the synthesis to the target, but its participation in substitution could not be discounted (to form a polypeptide polymer) so this possible complication was flagged.

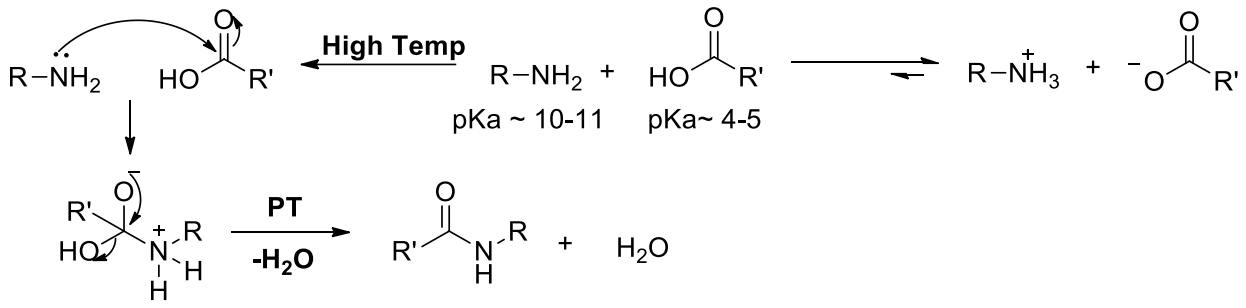


Scheme 3.4: The initial key target scaffold.

3.2.4 Amide Bond Formation

Following the synthesis of the 2-arylbenzimidazole key intermediate, formation of the amide bond, for which several methods are available, was the next and final step needed to reach the target compounds. The amide bond functionality is found in a whole host of molecules, most importantly as the peptide linkage in proteins, which are essential for life as they play important roles in a multitude of biological processes such as transport, immune protection and enzymatic catalysis. A survey of the Comprehensive Medicinal Chemistry database revealed that amide bonds occur in more than 25% of all known drugs, which is not surprising as they are known to be stable, neutral and have hydrogen bonding capabilities.[146,147] Amide bond formation between an amine and acid is formally called a condensation reaction

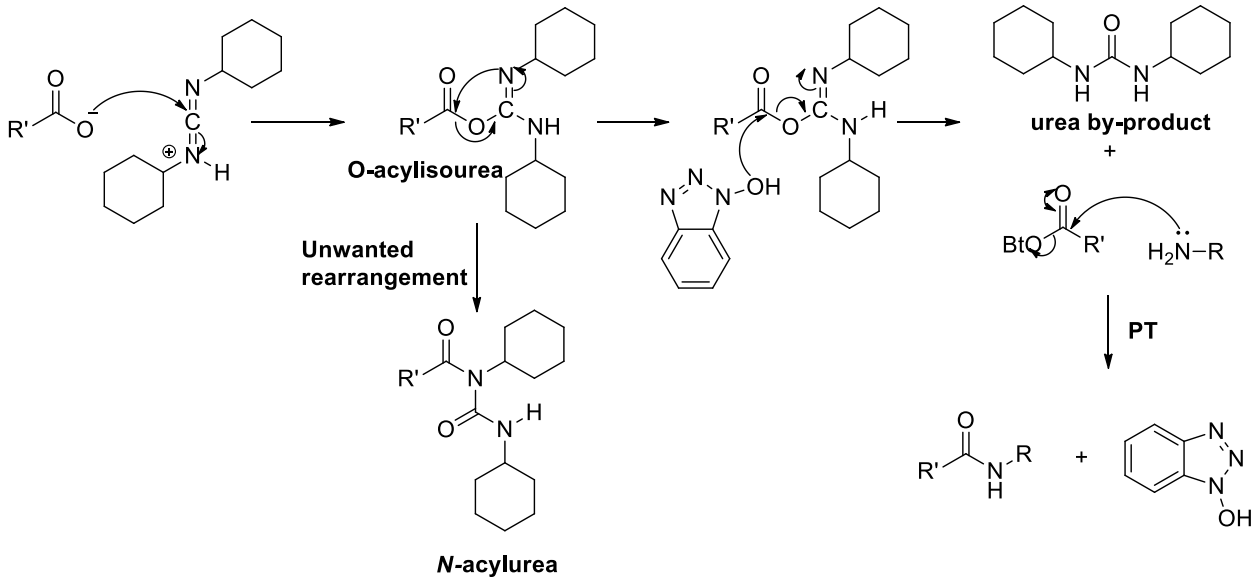
because of the expulsion of water during the reaction, which needs to be removed for the reaction equilibrium to proceed favourably to product. Indeed, mixing of the amine and acid forms a stable salt shown in Scheme 3.5, which prevents any condensation reaction from occurring. However, heating the reaction to temperatures greater than 160 °C allows the amine to override the inherent weak electrophilicity of the acid for the addition step of the reaction to occur, but such a high temperature is invariably incompatible with other functionalities. Reducing the activation energy for nucleophilic attack to occur requires an activation in the form of converting the hydroxyl group of the acid into a suitable leaving group.[147]



Scheme 3.5: Thermodynamics of amide bond formation.

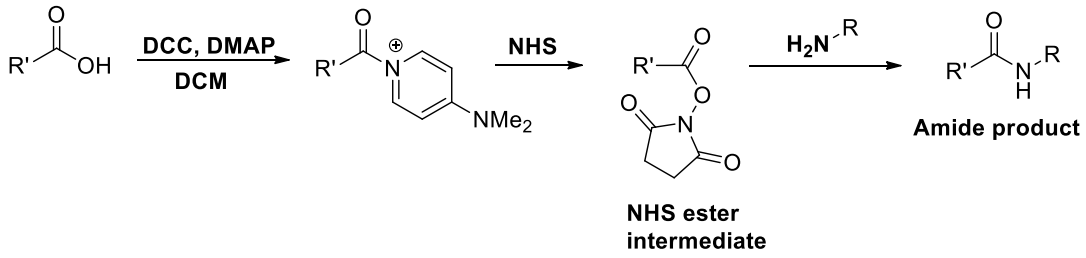
The literature contains a plethora of methodologies for achieving such activation, and these can be broken down into four main categories: namely carbodiimides, *N*-hydroxysuccinimide (NHS) esters, mixed anhydrides and acyl chlorides.[147] The carbodiimide methodology is mainly used in the synthesis of peptides as the other methodologies pose a danger of hydrolysis, racemisation and the cleavage of protecting groups. The latter will therefore be discussed first as an example to demonstrate the lowering of the activation energy for the amide formation reaction. The first carbodiimide mentioned in the literature was *N*,*N'*-dicyclohexylcarbodiimide (DCC), which in view of difficulties in removing its reaction by-product dicyclohexylurea during purification, has been largely replaced by second-generation diimides such as *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide (EDC), diisopropyl carbodiimide (DIC) and (1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate) (HATU). However, all of these follow a similar mechanism[147–149] and 1-hydroxybenzotriazole (HOBt) is invariably added to limit *O*- to *N*-acyl transfer as well as limit racemisation in the case of α-amino acids. Firstly, the carboxylic acid reacts with the carbodiimide in an acid catalysed process (via transfer from the carboxylic

acid) to form an *O*-acylisourea mixed anhydride intermediate that lowers the activation energy needed for subsequent nucleophilic attack. This intermediate first reacts with the oxygen of HOBt to form the Bt ester, which is then rapidly substituted by the amine via a S_NAc mechanism to furnish the product (Scheme 3.6).[147] The reaction can be carried out in a range of solvents, including water, dimethylformamide (DMF) and pyridine, in which EDC is probably the most commonly used reagent as it gives a water soluble urea by-product.[147]



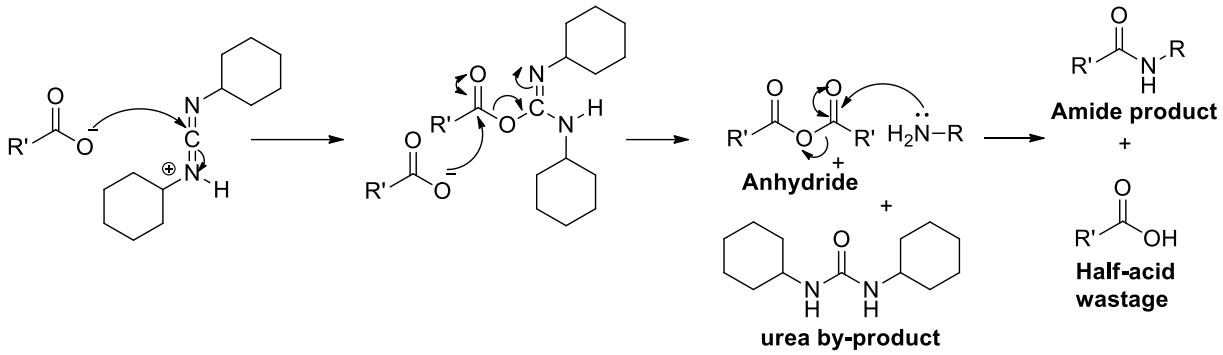
Scheme 3.6: Amide coupling reaction mechanism using DDC and HOBt.

The second methodology listed involves an esterification reaction that was first described by Steglich in 1978 involving 4-dimethylaminopyridine (DMAP) as an acyl-transfer catalyst.[150] This involves the formation of an NHS ester using DCC coupling described before. The NHS ester intermediate is then reacted with an amine to form the desired amide bond found in Scheme 3.7.[150]



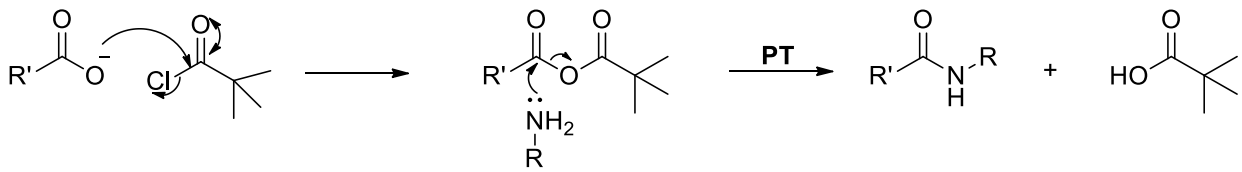
Scheme 3.7: The use of Steglich's esterification reaction to form an amide bond.

The third methodology listed involves the use of anhydrides, which are known to readily react with nucleophiles such as amines. These anhydrides are formed by heating the acid directly or via reaction of two equivalents of acid with one equivalent of DCC. This is then followed by a S_NAc reaction with an amine to furnish the amide bond shown in Scheme 3.8.



Scheme 3.8: Formation of an amide bond via the anhydride methodology.

However, only the half acid is used and thus large wastage occurs.[147] To alleviate this problem a mixed anhydride can be used where the second carboxylic moiety is cheap and easy to couple via its acid chloride.[151] An example is the pivalolyl group in which the bulk of the *t*-butyl group ensures a regioselective coupling (Scheme 3.9).[147]

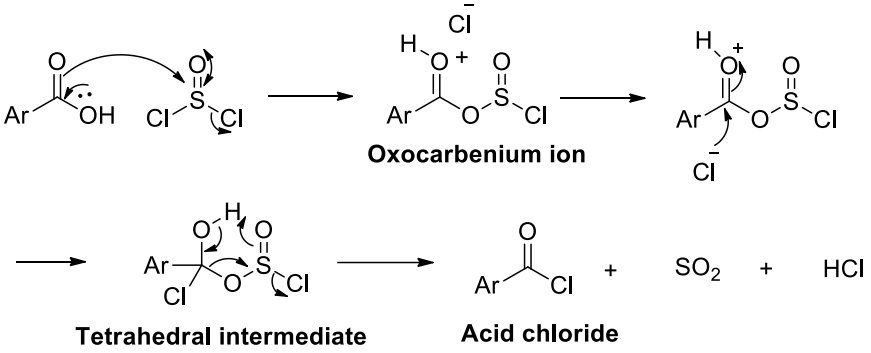


Scheme 3.9: Mixed anhydride reaction mechanism using pivaloyl chloride.

Acylation using an acid chloride is still one of the simplest and effective methodologies so long as one is not worried about hydrolysis, racemisation or environmental concerns (small scale). Many acid chlorides are commercially available, and if not, can be prepared from the acid using one of many chlorinating agents. These include oxalyl chloride, phosphorus pentachloride, phosphorus trichloride, phosphorus oxychloride or thionyl chloride.[147,152] Notably, though, care should be taken when working with these reagents as they release toxic hydrogen chloride (HCl) gas.

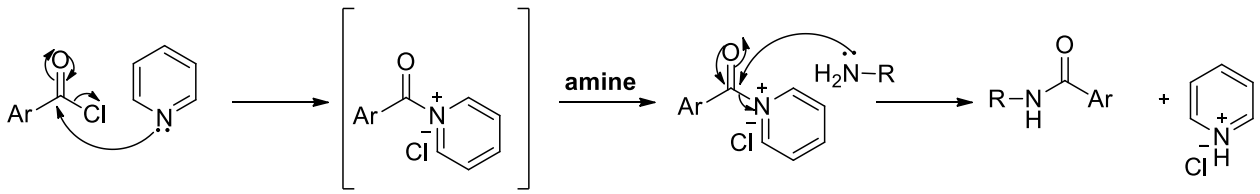
The mechanism for the formation of an acid chloride using thionyl chloride is well known throughout the literature. Initially, the reaction begins with the reaction of the acid with thionyl chloride to form a chlorosulfite ester in the form of a protonated carbonyl group

(oxocarbenium ion). The latter activates the addition of chloride to generate the tetrahedral intermediate that is believed to break down in an entropically favoured, concerted fashion to expel SO_2 and HCl gas with formation of the acid chloride. Overall, the conversion constitutes a S_NAc reaction as shown in Scheme 3.10.[147,152]



Scheme 3.10: Acid chloride formation mechanism.

Once the acid chloride is formed, the appropriate amine can be coupled in yet another S_NAc reaction to form the amide bond, in which a base is required to neutralise the HCl produced. Mild, non-nucleophilic bases are required here to avoid nucleophilic attack on the chloride by the base (eg. hydroxide does not qualify).[152] In this regard, amine bases are popular such as triethylamine (Et_3N), Hünig's base (*i*-Pr_2NEt) or pyridine. Of these, pyridine is the most favourable as it acts as an acyl-transfer agent, which increases the rate of the reaction. Together with its ability to improve the solubility of substrates, pyridine makes for a popular choice of solvent in these types of reactions.[147] The transfer aspect of pyridine is interesting (see Scheme 3.7) and is represented below in Scheme 3.11.



Scheme 3.11: Acetylation reaction mechanism using pyridine.

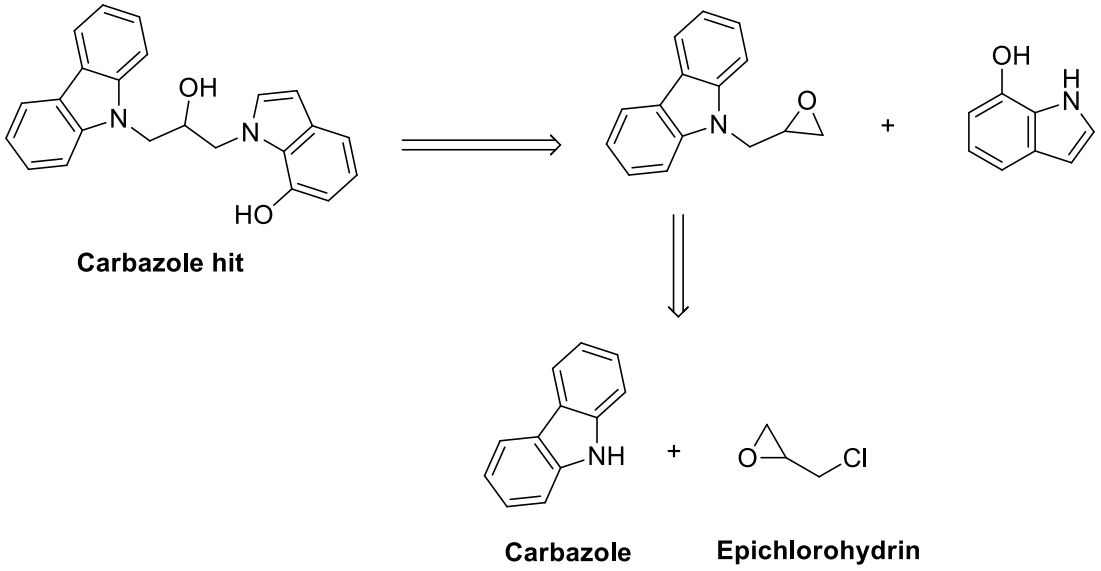
Pyridine undergoes a S_NAc reaction with the acid chloride faster than the primary amine owing to its nitrogen's strong Lewis basic and soft character. This results in an *N*-acyl-pyridinium salt, which is more electrophilic than the acid chloride. Thereafter, a fast amine attack ensues with regeneration of pyridine, which, following proton transfer affords the amide and pyridinium hydrochloride as a by-product.[147,152] With so many methodologies

available the suitability of reagents, conditions, the yield, formation of unwanted side products, purification and the selectivity of the reagents needed to be considered.[147] Ultimately, the acid chloride methodology was decided upon for the purpose of this thesis.

3.3 Scaffold 2: Carbazole Indoles

3.3.1 Retrosynthetic Analysis of the Carbazole Indole Compound

Similar to the benzimidazole case previously, the proposed carbazole hit analogue could be disconnected back to 7-hydroxyindole and an epoxide, which may then be disconnected back to epichlorohydrin and carbazole as shown in Scheme 3.12. The forward direction therefore consisted of an S_N2 reaction between the carbazole and epichlorohydrin followed by a regioselective epoxide opening to realize the target compounds. The IC_{50} values measured for the original hit compound were obtained using the racemate, therefore separate enantiomers of the hit analogue were synthesized and evaluated. As with the benzimidazole hit compound, various aspects of the carbazole moiety will now be discussed.



Scheme 3.12: Retrosynthetic analysis for the carbazole indole hit compound.

3.3.2 Carbazole Subunit

Carbazoles are important heterocycles in medicinal chemistry, consisting of a 5-membered pyrrole core fused symmetrically with two benzene rings,[124] making it an indole ring with a benzene ring fused onto the 2,3-positions. The classical method for the preparation of the carbazole moiety is through the dehydration of 1,2,3,4-tetrahydrocarbazoles prepared by the

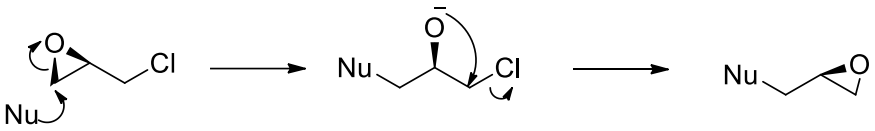
Fischer-Borsche or the Graebe-Ullmann synthesizes. However, a more recent methodology involves the use of transition metal catalysis (iron, palladium or molybdenum).[118] For the purposes of this thesis all carbazole reagents were purchased, and the term carbazole refers to 9*H*-carbazole moiety, with its reactivity based on the nucleophilic nature of the deprotonated nitrogen anion.[118,124]

3.3.3 Indole Subunit

Indoles are an important bicyclic heterocycle that consist of a pyrrole ring fused to a benzene ring at the α, β-positions. They occur in proteins as the amino acid tryptophan and in indole alkaloids such as LSD. The indole ring system is a widely encountered pharmacophore in medicinal chemistry.[153] The prevalence of indole chemistry can be seen with nine different named reactions for the synthesis of this pharmacophore in which the Fischer indole synthesis is the most famous.[154] The reactivity of this substrate in context again lies in the nucleophilic nature of the deprotonated nitrogen anion and for the purpose of this thesis, the indole ring systems were purchased.[153,154]

3.3.4 Epoxidation using Epichlorohydrin

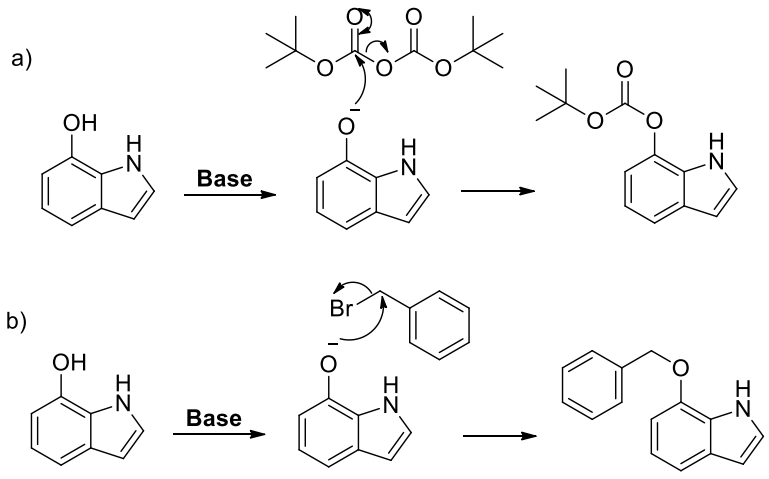
Epoxides are generally formed by oxidising an alkene using a peroxy acid, most commonly *m*-chloroperoxybenzoic acid (*m*-CPBA), but for attachment of a three-carbon chain epoxide to a secondary amine, epichlorohydrin is the more effective choice.[152] Epichlorohydrin can be purchased as a racemic mixture or as chirally pure enantiomers, therefore it is possible to form the different enantiomer target products. However, it is considered to be a carcinogen, so extreme caution must be taken during use.[155] Interestingly, based on the reaction of unsymmetrical (substituted) epoxides during substitution, nucleophilic attack occurs at the epoxide carbon rather than directly on the carbon bearing the chloride. With strong, relatively soft nucleophiles the epoxide opening reaction involves an S_N2 mechanism in which the nucleophile attacks the least hindered side of the epoxide, opening it up and forming an oxygen anion. This anion attacks the carbon adjacent to the chlorine, eliminating a chloride and forming the desired product with regeneration of the epoxide and with retention of configuration of the C-O bond at the stereogenic centre, as shown in Scheme 3.13.[152]



Scheme 3.13: Epichlorohydrin substitution mechanism.

3.3.5 Protection of the Alcohol

After forming the intermediate compound, a methodology for regioselective alkylation (using the epoxide) on the amino nitrogen rather than the phenolic hydroxyl oxygen of 7-hydroxyindole needed to be established. Base catalysis to afford the indole *N*-anion was envisaged as necessary for this reaction in view of the low nucleophilicity of the indole nitrogen, which meant that protection of the relatively acidic hydroxyl group was required. The chemical literature contains a large number of reagents for alcohol protection, but in context two groups stood out as likely candidates owing to their ease of introduction and removal.[152] These were a benzyl ether and a *N-tert*-butoxycarbonyl (Boc) carbonate, for which relevant mechanisms are shown in Schemes 3.14a and 3.14b. Both involve different types of substitution (S_N2 and S_NAc respectively) mechanisms involving the oxide anion of the indole.[152] The benzyl protection methodology was chosen for the purposes of this thesis due to the stability of this group in basic conditions and at high temperatures.

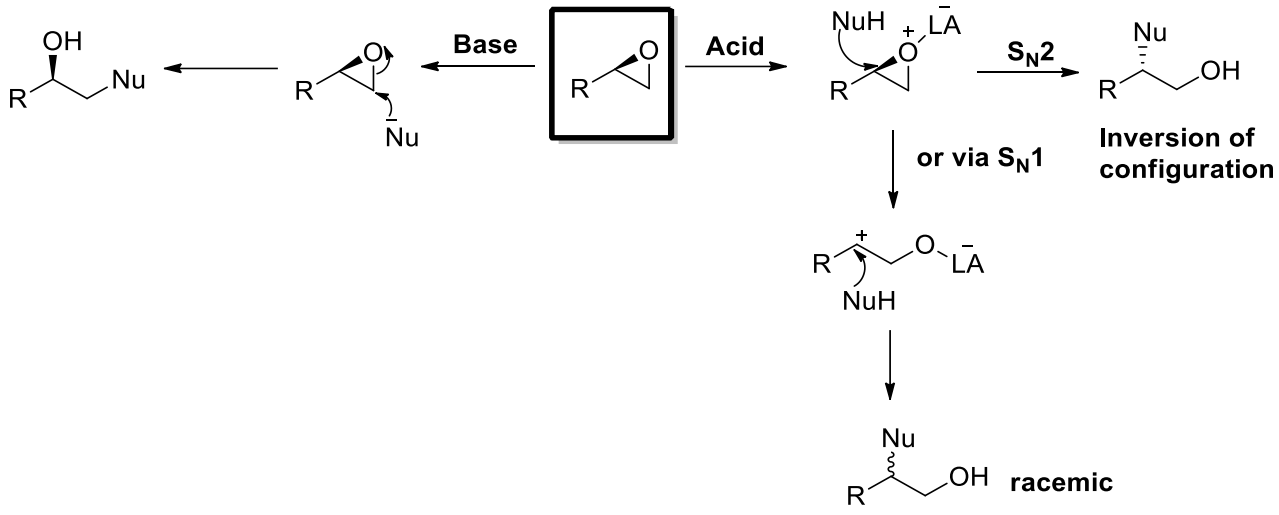


Scheme 3.14: a) Boc protection mechanism, b) Benzyl protection mechanism.

3.3.6 Epoxide Opening

With the protection of the alcohol in hand, various aspects of the epoxide opening needed to be considered. While nucleophilic attack at the chiral carbon (more substituted) of a chiral

epoxide of a terminal alkene is known to be promoted under Lewis acidic conditions in view of the secondary carbocation character in the transition state, the requirement here for attack at the less-hindered carbon meant that base catalysis via an S_N2 manifold was required (refer to Scheme 3.15 for both mechanisms).[152] The rate of an S_N2 reaction decreases with the steric bulk of the nucleophile, which suggested that a bulky nucleophile was not ideal.[152] Additionally, it is important to note that such an S_N2 reaction (at the least hindered epoxide carbon) would retain stereogenicity at the C-O bond of the chiral centre, which was an important aspect of the synthetic design.



Scheme 3.15: Acid and base mediated epoxide opening mechanisms where LA represents lewis acid.

3.4 Synthetic Goals

The goals of this chapter were to:

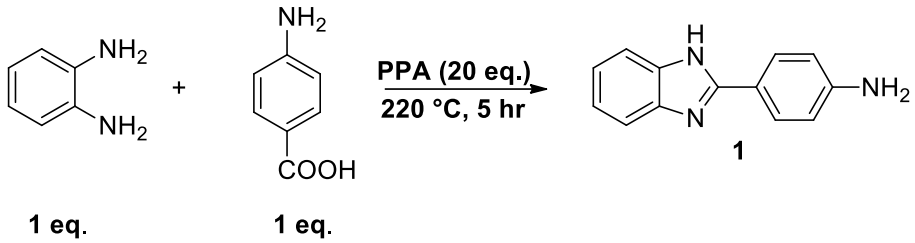
- Synthesize and characterize the eighteen benzimidazole compounds identified by the Bayesian statistics.
- Synthesize and characterize the seven benzimidazole substructures from the Bayesian study.
- Synthesize and characterize the seven carbazole indole compounds from the Bayesian study.

3.5 Results and Discussion

3.5.1 Scaffold 1: Benzimidazoles

3.5.1.1 Synthesis of the Benzimidazole Intermediate

Formation of **1** was carried out on a 40 mmols scale using the reaction conditions detailed in Scheme 3.16 following the previously established methodology of Chua *et al.*[156] After reaction completion, which was verified by TLC analysis, $NaHCO_3$ was used to neutralize the acid in a mini work-up, the mixture was transferred into a beaker quickly using warm water, followed by several washes. This mixture was cooled using an ice bath and slowly neutralized using $NaHCO_3$ so as to minimise reaction vigour (CO_2 evolution). It was imperative that the reaction was not overly basified so as to avoid substantial losses of deprotonated product into the water. Column chromatography was not required as recrystallization of the solid obtained (via filtration) using methanol was used to achieve a yield of 61% with an HPLC purity of ≥ 95%.



Scheme 3.16: Reaction conditions for the formation of the benzimidazole product (**1**).

The synthesis of the benzimidazole product was confirmed by its ^1H nuclear magnetic resonance (NMR) spectrum using deuterated dimethyl sulfoxide (DMSO-d_6) as a solvent to ensure visualisation of the primary amine peak as shown in Figure 3.1. As expected, two doublets typical of a 1,4-disubstituted system, accounted for H-11, H-12, H-14 and H-15, while two multiplets accounted for the benzimidazole ring system. One is reminded here that rapid N-H tautomerism renders the benzimidazole A-ring protons symmetrically equivalent in ^1H NMR. A two-proton broad singlet at 5.54 ppm confirmed the presence of the primary amino group, and this signal could be used as a diagnostic marker for this moiety. Similarly, the ^{13}C NMR spectrum showed eight distinct signals as expected, four for each aromatic ring based on the above mentioned symmetry grounds. Definitive assignment of the ^{13}C NMR spectrum resonances could not be ascertained for compound **1**.

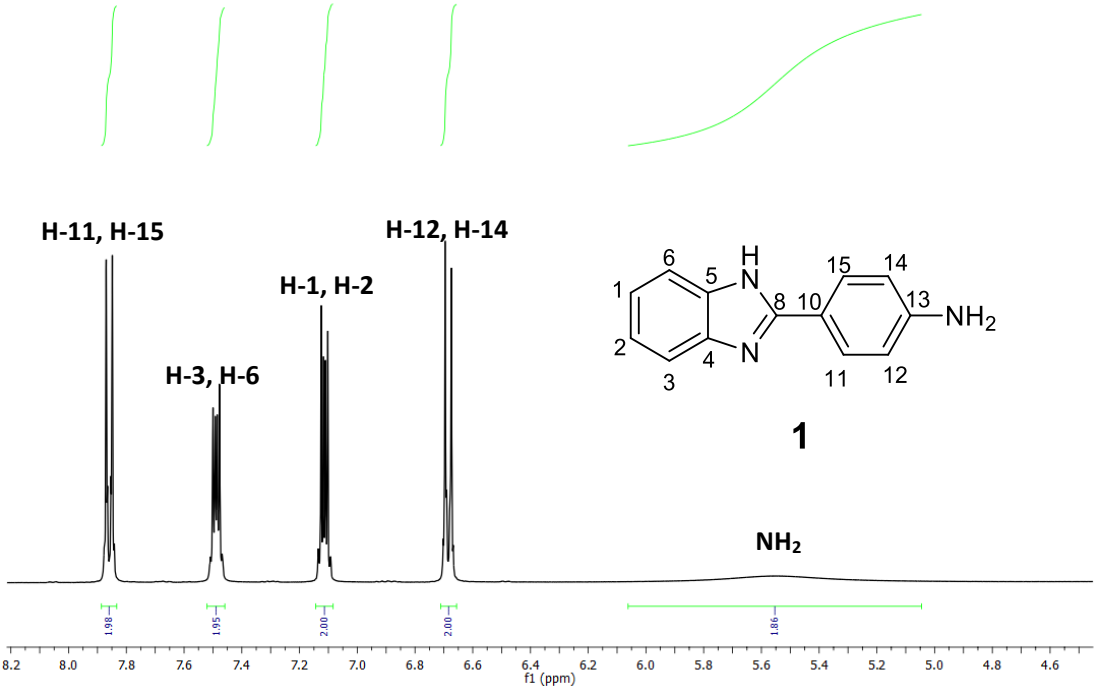
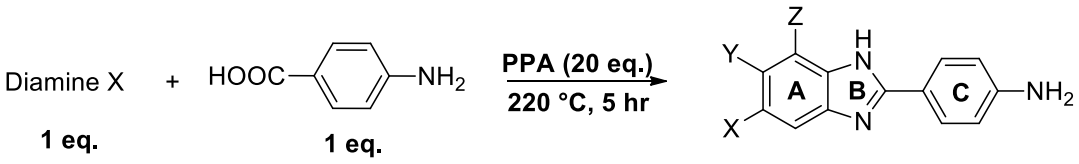


Figure 3.1: The ^{1}H NMR spectrum of the benzimidazole intermediate (**1**) in DMSO-d_6.

Following the successful formation of **1**, attempts were then made to synthesize a variety of derivatives of **1** using the same methodology described but using various substituted *o*-phenylenediamine derivatives (**1a**-**1g**), shown in Scheme 3.17, so as to functionalise the A-ring.



Diamine X
1a: 4-Nitro-*o*-phenylenediamine
1b: 3-Nitro-*o*-phenylenediamine
1c: 4-Chloro-*o*-phenylenediamine
1d: 4,5-Dichloro-*o*-phenylenediamine
1e: 4-Fluoro-*o*-phenylenediamine
1f: 4-Methyl-*o*-phenylenediamine
1g: 3-Methyl-*o*-phenylenediamine

Products
1h: X=H, Y= NO_2, Z= H
1i: X=H, Y= H, Z= NO_2
1j: X=H, Y= Cl, Z= H
1k: X=Cl, Y= Cl, Z= H
1l: X=H, Y= F, Z= H
1m: X=H, Y= Me, Z= H
1n: X=H, Y= H, Z= Me

Scheme 3.17: The reaction conditions for the attempt at synthesizing compound **1** derivatives.

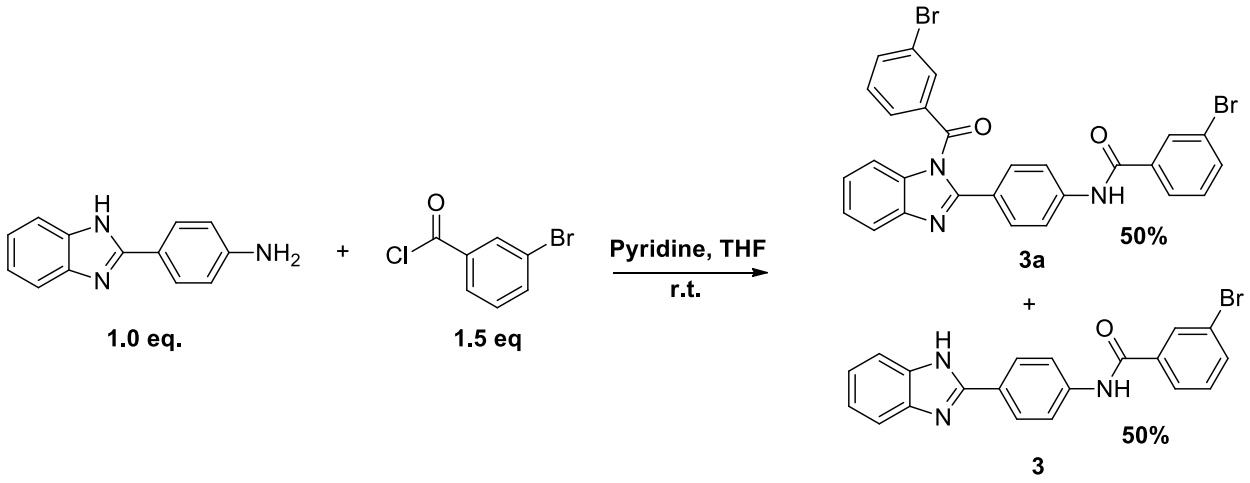
All of the substituted *o*-phenylenediamine reactants (**1a**-**1g**) degraded (formed a black precipitate) during the reaction, presumably due to the high temperatures (220 °C) required for these reactions to proceed in the presence of PPA (strong acid). These reactants (**1a**-**1g**) were found to be stable in PPA at room temperature and so a temperature study was carried out on all of the diamine reactants **1a**-**1g** using the reaction conditions detailed in Scheme 3.17 (i.e. with the PPA) to determine the optimum temperature for reaction. The melting point of each diamine (**1a**-**1g**) was taken as a starting temperature for initial reaction. The reactions were analysed using TLC every hour to determine reaction progress and the degradation point of each compound. If no reaction had occurred the temperature was increased by 10 °C. This was carried out until the compound degraded or a maximum temperature of 220 °C was reached. All compounds fully degraded before they reached 220 °C, with no distinct product visualised. Therefore, it was concluded that the formation of condensation products (**1h**-**1n**) was not possible using these conditions, and thus this route was abandoned.

3.5.1.2 Acid Chloride and Amide Bond Formation Results

After the disappointing results from the temperature study, attention was redirected to the C-ring by preparing a library of amides of intermediate **1**. However, it was first imperative to consider access to the acid chlorides. This was achieved either by reacting the relevant acid with thionyl chloride or accessing the acid chloride commercially directly from commercial sources. Using 3-bromobenzoic acid as a pilot study, the acid chloride was formed by refluxing the acid in thionyl chloride overnight to ensure complete conversion. TLC could be used to monitor the progress of the reaction by quenching a drop of the mixture with MeOH to form the methyl ester. Progress could be monitored comparing residual acid versus the ester due to the polarity difference on the TLC plate. Once the esterification was complete, the excess thionyl chloride was reduced under pressure and the newly formed acid chloride was used within an hour with no further purification.

Initial reaction of 3-bromobenzyl chloride (1.5 eq.) and **1** (1.0 eq.) was performed by stirring the mixture in anhydrous tetrahydrofuran (THF), with excess pyridine, at room temperature. However, formation of an unwanted side product was identified, whose ^{1}H NMR and ^{13}C NMR spectra both showed the presence of two amide functionalities rather than the desired one,

as well as an additional aromatic ring system. This resulted in the structure being assigned as **3a**, as shown in Scheme 3.18**.**



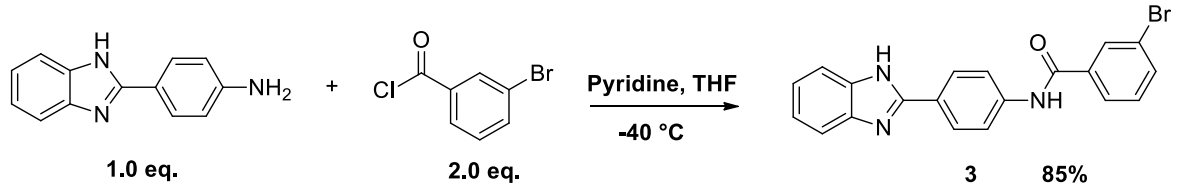
Scheme 3.18: Trial acid chloride coupling reaction.

These results reflected that the reaction was driven by the electrophilicity of the *N*-acylpyridinium ion, which could promote acylation at both types of nitrogen (the free amino and the heteroaromatic N). Evidently, the soft but nucleophilic amidine of the imidazole ring of the benzimidazole is reactive enough to be acylated. However, TLC analysis suggested that formation of compound **3** via pathway 1 was faster in that compound **3** formed first.

Using this reasoning the number of equivalents of 3-bromobenzyl chloride was thus lowered to one, still using room temperature as the reaction temperature. However, this now resulted in an incomplete conversion in which a portion of diacylated product was still visible on TLC. Column chromatography separation of **1** from any of the products (**3** and **3a**) was not fully possible due to the small retention factor (Rf) difference between the compounds as well as their propensity to 'drag' on the column.

Despite these results, it was decided to next examine the effect of temperature on the reaction. This was achieved by varying the reaction temperature of four separate reactions, using 1.5 eq. of 3-bromobenzoyl chloride and 1.0 eq. of **1**. The first reaction at room temperature yielded the same results as before with a 1:1 ratio of **3a** to **3** visualised on TLC. The reaction at 0 °C achieved the same result with a 1:1 mixture of **3a** and **3** visualised on TLC.

Using the same methodology as before, but further lowering the temperature to -78 °C, yielded no reaction as the pyridine froze. The last attempt at -40 °C was successful, with formation of only compound **3** visualised on TLC. However, residual **1** was also visualised on TLC which was also undesirable. In order to mitigate this problem, two further reactions were attempted increasing the original 1.5 eq. of 3-bromobenzoyl chloride used in the reaction. An increase to 1.8 eq. still yielded the same results as with the 1.5 eq. Finally, increasing the equivalents of 3-bromobenzoyl chloride to 2.0 resulted in full conversion with no formation of **3a** visualised via TLC**.** The optimised reaction conditions are shown in Scheme 3.19.



Scheme 19: Optimised reaction conditions for full conversion to compound **3**.

The product was too polar for conventional extraction, therefore the bulk of the pyridine was azeotropically removed with toluene at 60 °C on a rotary evaporator. The remaining residual pyridine was reduced under pressure before subjecting the crude mixture to column chromatography using MeOH and DCM mixtures. It was assumed that the HCl produced in the reaction stayed on the column as pyridinium hydrochloride. Further purification for biological evaluation purposes was achieved by recrystallizing product **3** from methanol and water mixtures to constant melting point of 324 - 325 °C.

The ^{1}H NMR spectrum of **3** is shown in Figure 3.2. This showed two multiplets at 7.20 ppm and 7.63-7.79 ppm accounting for 2H and 3H respectively, the latter corresponding to the A-ring protons and H-21. Following this, a signal at 7.81 ppm accounted for 1H with a *J* value of 7.8 Hz which was assigned to H-22. Two other multiplets were seen at 7.98 and 8.19 ppm, integrating for 3H each and were assigned to H-12, H-14 and H-20 and H-11, H-15 and H-24 respectively. The diagnostic signal for the amide proton H-16 was found at 10.54 ppm integrating for 1H. Finally, the assignment of H-9 at 12.80 ppm implied no formation of **3a**.

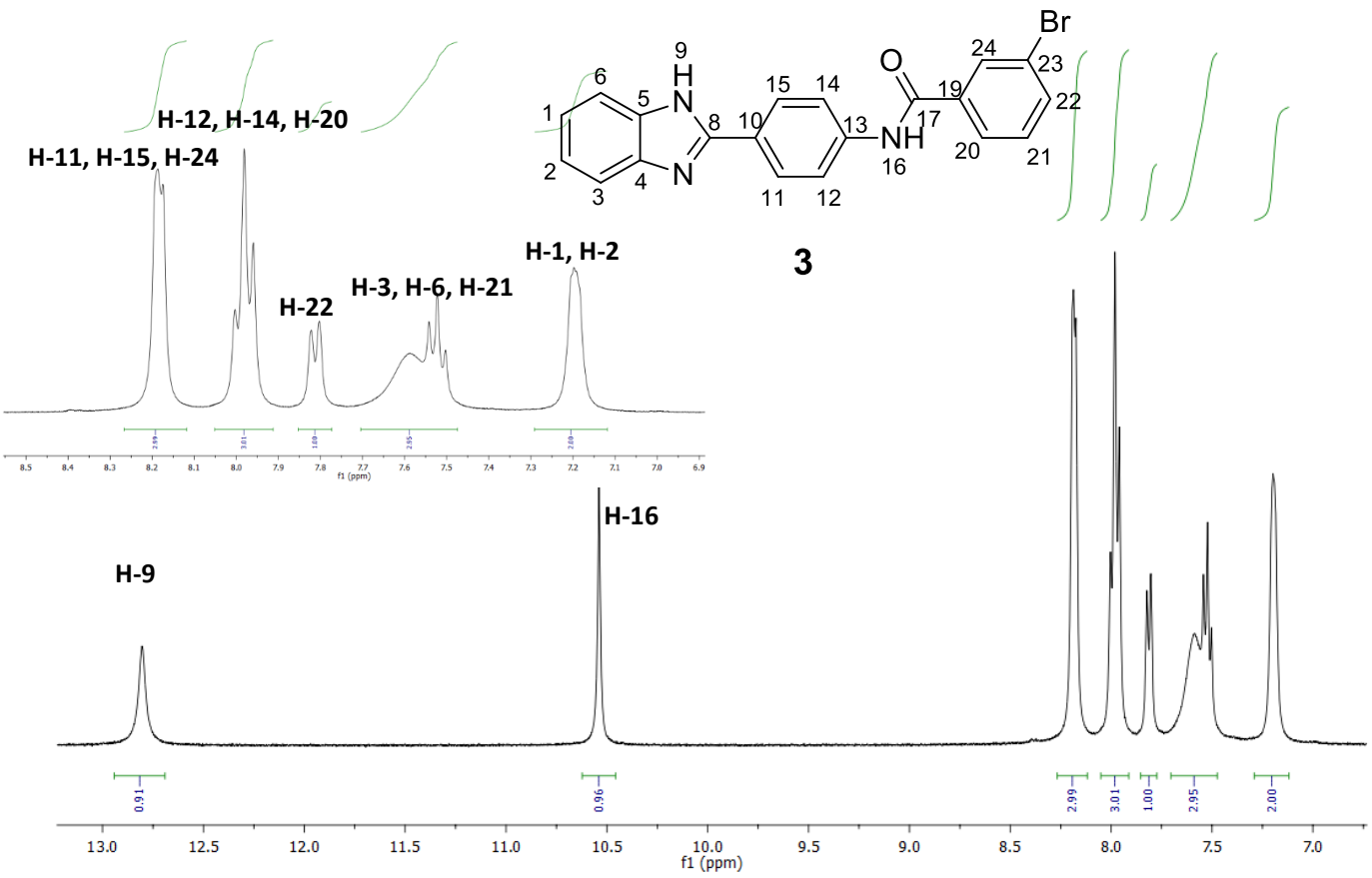


Figure 3.2: The ^{1}H NMR spectrum of compound **3** in DMSO-d_6.

The ^{13}C NMR spectrum of **3** is shown in Figure 3.3, where the carbonyl carbon signal at 164.1 ppm was diagnostic of the formation of the amide bond. Overall, eighteen resonances (singlets) were observed with a 7:4:6 split (benzimidazole: C: D) plus the carbonyl, revealing that desymmetrization of the benzimidazole A ring system was observed implying a slow benzimidazole H-tautomerism on the NMR time scale. A reminder of this tautomerism is shown in Scheme 3.20 to illustrate this.

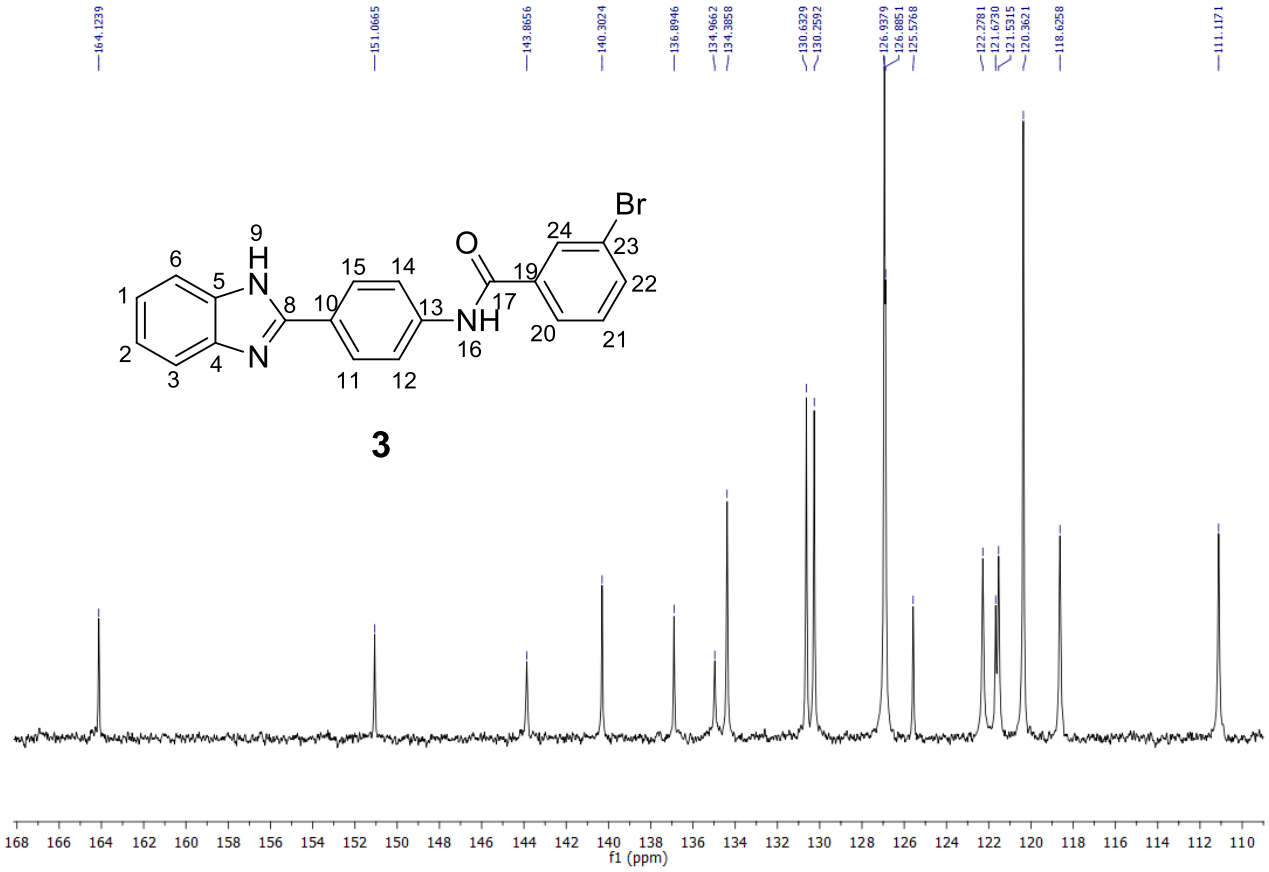
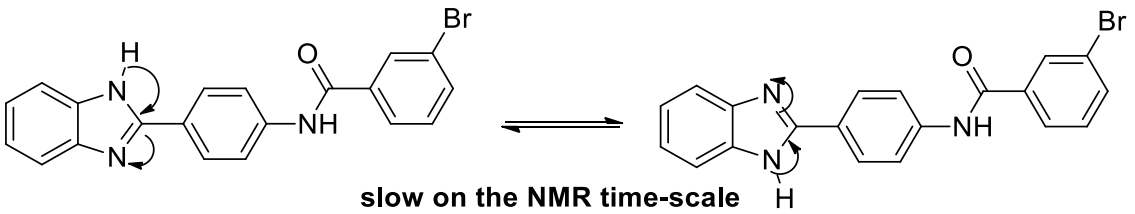
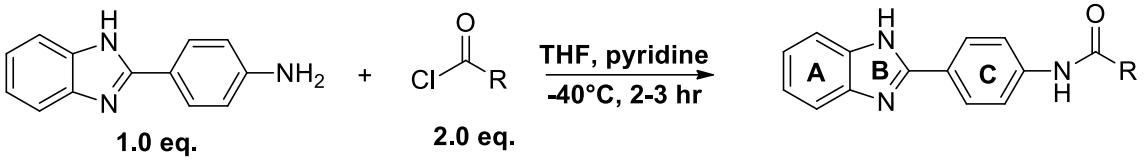


Figure 3.3: The ^{13}C NMR spectrum of compound **3** in DMSO-d_6.



Scheme 3.20: Tautomerism of the benzimidazole ring system of compound **3**.

This overall methodology was followed for all the derivatives for this study and is depicted in Scheme 3.21 except for the recrystallizing solvent used. The reaction times differed slightly depending on the acid chloride used, although no correlation could be found between the reaction times and the nature of the substituents on the acid chloride based on its electron-withdrawing or donating groups. Rather, the insolubility of the acid chloride emerged as the main determinant, with the insoluble acid chlorides (**5**, **9, 12** and **16**) producing the longer reaction times, which changed from two to three hours.



Phenyl Derivatives
2: R = Phenyl
3: R = *m*-Bromophenyl
4: R= *m*-Trifluoromethylphenyl
6: R= *p*-*tert* Butylphenyl
7: R = 2,4-Dichlorophenyl
8: R = *o*-Methoxyphenyl
10: R= *o*-Nitrophenyl
11:R = 3,5-Dinitrophenyl
13: R = *o*-Chlorophenyl
14: R = *m*-Chlorophenyl
15: R = *p*-Chlorophenyl

Pyridyl Derivatives
5: R= 3-Bromopyridyl
9: R= 3-Trifluoromethylpyridyl
12: R = 2,6-Dichloropyridyl
16: R = 3-Pyridyl

5-Membered Ring Derivatives
17: R = Thiophenyl
18: R = Furanyl
19: R = Pyrrole

Scheme 3.21: Synthesized benzimidazole derivatives.

3.5.1.3 NMR Analysis and Other Characterisation of the Benzimidazole Derivatives

The library depicted in Scheme 3.21 was characterized using ^{1}H NMR, ^{13}C NMR, correlation spectroscopy (COSY) and heteronuclear single-quantum correlation (HSQC) NMR spectroscopy. DMSO-d_6 was used as a solvent for all of the NMR experiments, so as to solubilize the polar compounds and to visualize the amine peaks. In spite of the abundance of aromatic signals, a significant amount of information could be extracted from the NMR data. Firstly, the ^{1}H NMR spectra for all eighteen compounds were similar, showing the benzimidazole core containing two multiplets at ≈ 7.2-7.3 ppm and ≈ 7.6-7.7 ppm, respectively and were assigned as the benzimidazole ring system protons. While the two doublets at ≈ 8.0-8.1 ppm and ≈ 8.2-8.3 ppm, respectively were assigned to the 1,4-disubstituted phenyl system protons. Importantly, though, diagnostic ^{1}H NMR peaks for the formation of these compounds was provided by aromatic ring proton peaks from the acid chloride, as well as the disappearance of the primary amine at around 5.54 ppm in the starting material, which was replaced by a secondary amide peak found at ≈ 10 ppm, signifying the formation of the amide bond. Similar to compound **3**, the benzimidazole NH was visualized at ≈ 12.8 ppm signifying no formation of side products such as **3a**. These features can all be found in the ^{1}H NMR spectra for **5** and **12** shown in Figures 3.4 and 3.5.

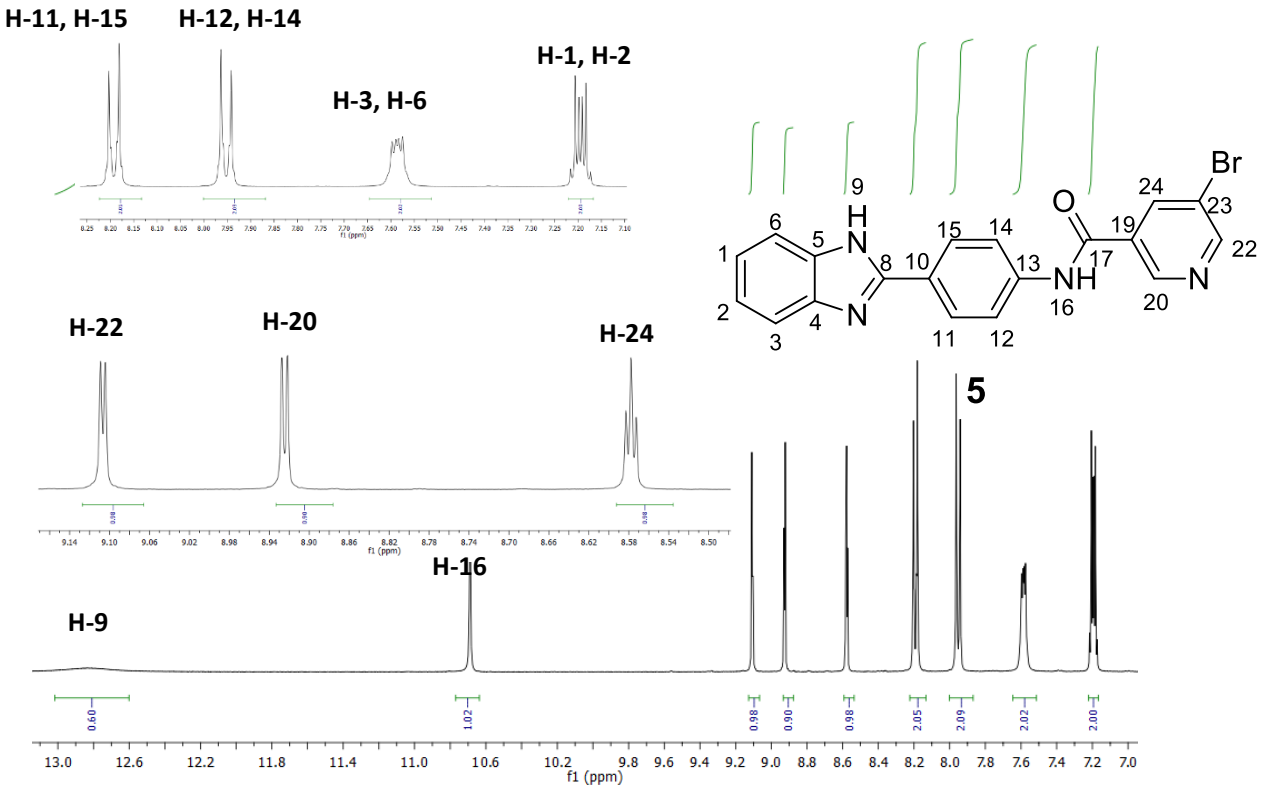


Figure 3.4: ^{1}H NMR spectrum of **5** in DMSO-d_6.

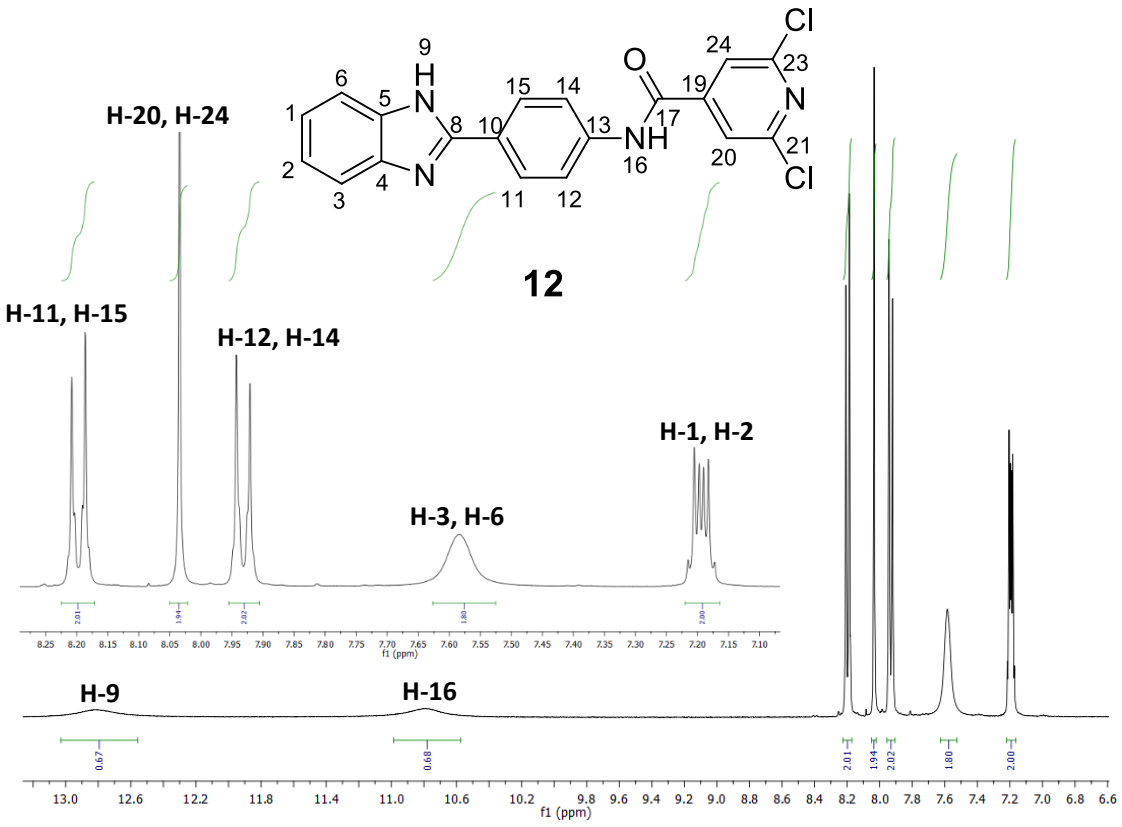


Figure 3.5: ^{1}H NMR spectrum of **12** in DMSO-d_6.

As expected, the ^{13}C NMR spectra for compounds **2**-**19** revealed the diagnostic carbonyl carbon peak for the amide at ≈ 161 ppm as well as new carbon peaks accounting for the addition of the new aromatic ring system. Interestingly, while the ^{13}C NMR spectra for compounds **3**, **4**, **5**, **7**, **10** and **12** showed seven carbon peaks for the benzimidazole ring system other compounds **2**, **6**, **8**, **9**, **11**, **13**, **14**, **15**, **16**, **17, 18** and **19** showed just four. As before, this was attributed to differences in the rates of NH-tautomerism, but a correlation between structure and number (as a relative tautomerism rate) could not be established. These subtleties are illustrated in Figures 3.6 and 3.7 for **14** (symmetrization) and **10** (desymmetrization) respectively.

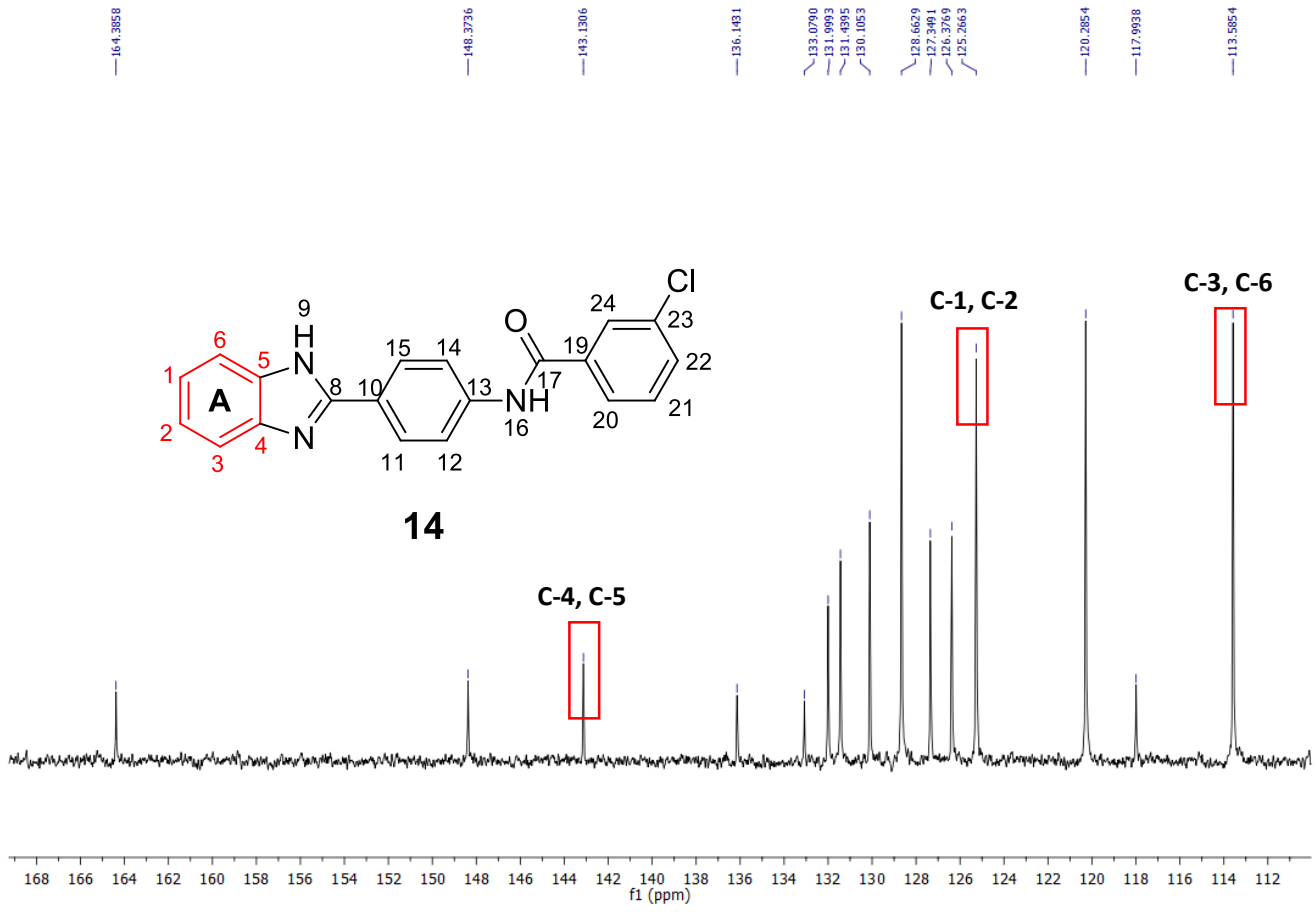


Figure 3.6: The ^{13}C NMR spectrum of **14** in DMSO-d_6, showing symmetrization of the A-ring carbon signals (shown in red).

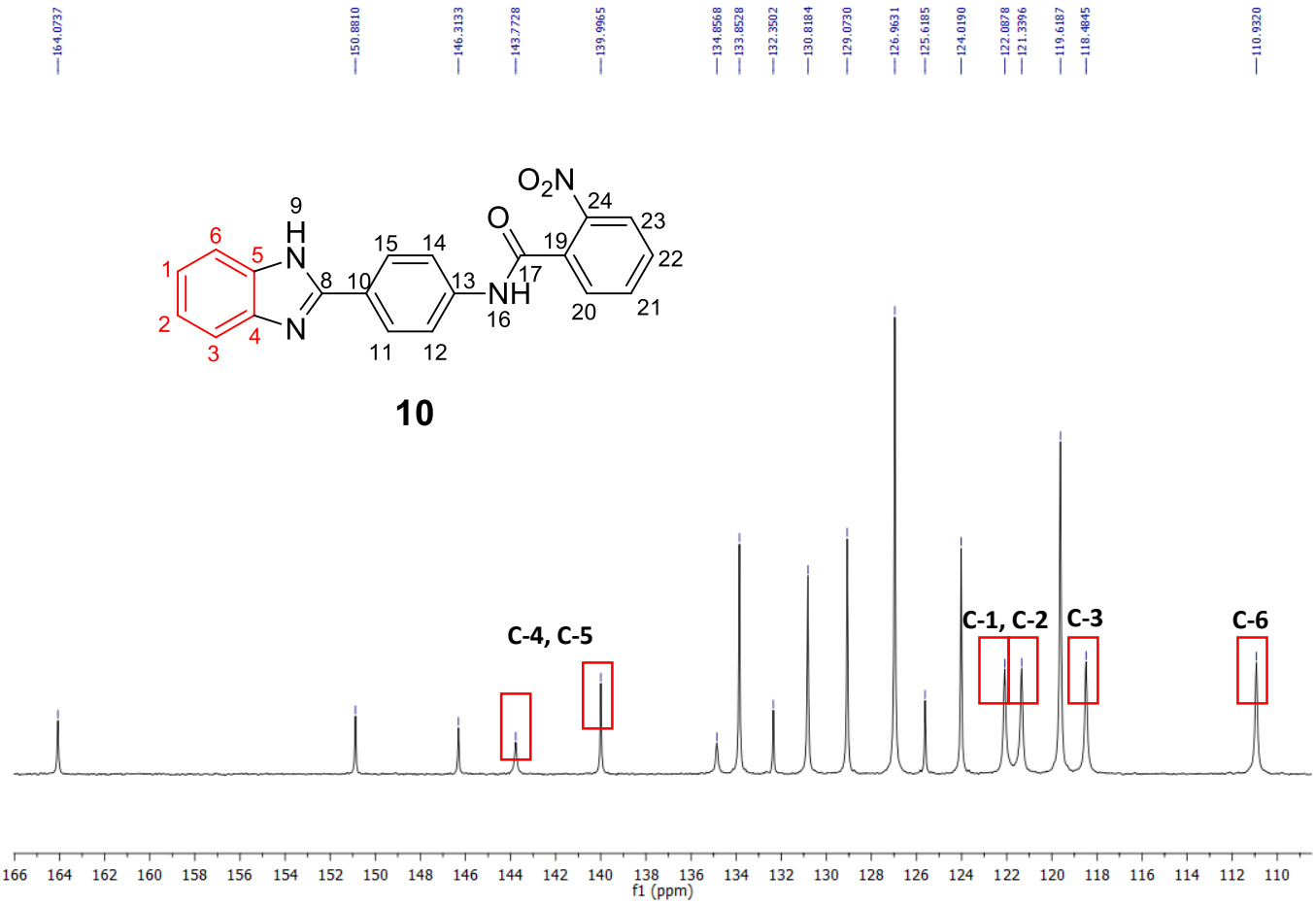
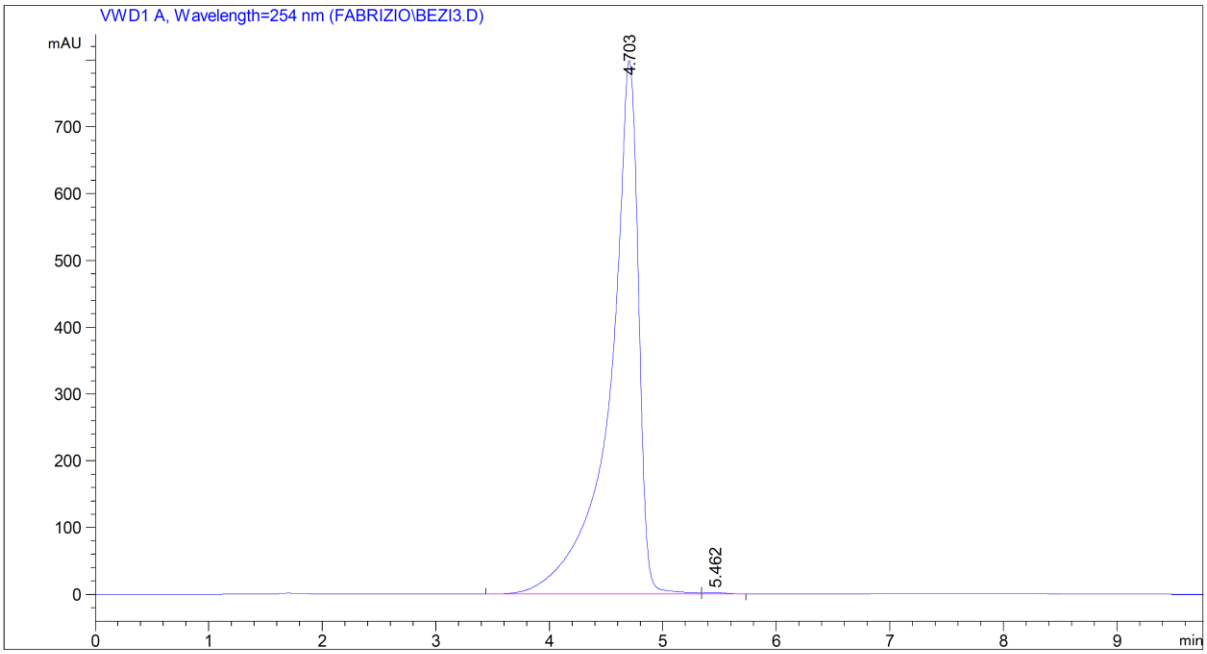


Figure 3.7: The ^{13}C NMR spectrum of **10** in DMSO-d_6, showing desymmetrization of the A-ring carbon signals (shown in red).

Other methods of characterization used for these compounds included high resolution mass spectrometry (HRMS), IR spectroscopy, melting point analysis (to within one degree of the previous melting point), and high-performance liquid chromatography (HPLC) to at least 95% purity with compound **3** shown as an example in Figure 3.8.

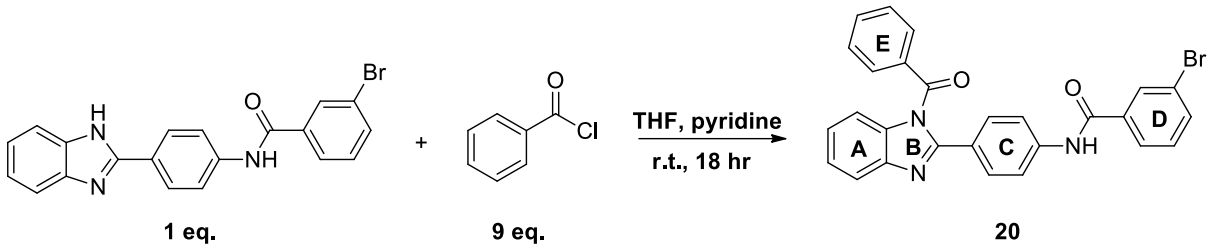



```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier:                   :      1.0000
Dilution:                     :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=254 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   4.703 BV    0.2436 1.39326e4   798.44977  99.7985
   2   5.462 VB    0.1923   28.13146    2.15429   0.2015

Totals :                  1.39607e4   800.60406
```

Figure 3.8: C18 HPLC trace of compound **3** at 254 nm.

Given the formation of **3a**, it was decided to include a similar compound in this part of the study to determine the effect of removing the NH hydrogen bond in the B-ring on the β-haematin and parasite growth inhibitory activities. Therefore, **20** was formed using compound **3** (1.0 eq.), THF, pyridine and benzoyl chloride (9.0 eq.) at room temperature for 18 hours, as shown in Scheme 3.22. Upon complete disappearance of **3**, monitored via TLC analysis, a standard work-up procedure was used involving acidifying the reaction mixture to pH 4 using HCl to remove most of the pyridine from the organic layer. **20** was then extracted using MeOH and chloroform mixtures. The crude mixture was then subjected to column

chromatography using MeOH and DCM mixtures to afford **20** in 39% yield, which was followed by recrystallization using acetone and water to yield a crystalline solid.



Scheme 3.22: Reaction conditions for the formation of compound **20.**

The ^{1}H NMR spectrum of product **20**, shown in Figure 3.9, showed the same proton signals as those of compound **3** but with the addition of five extra aromatic ring protons, accounting for the E-ring**.** The diagnostic carbonyl carbon peak for **20**, at 168.9 ppm could be seen in the ^{13}C spectrum (Figure 3.10). The ^{13}C NMR spectrum of **20** also showed the expected four additional carbon peaks accounting for the E-ring carbons as well as twenty three resonances overall, based on a fully desymmetrized A-ring due to benzimidazole *N*-substitution. Further confirmation of compound **20** was given by virtue of a correct HRMS HRMS-ES^{+} Calculated: 496.0661 $[M+H]^{+}$ for $C_{27}H_{19}BrN_3O_2$, Observed: 496.0654.

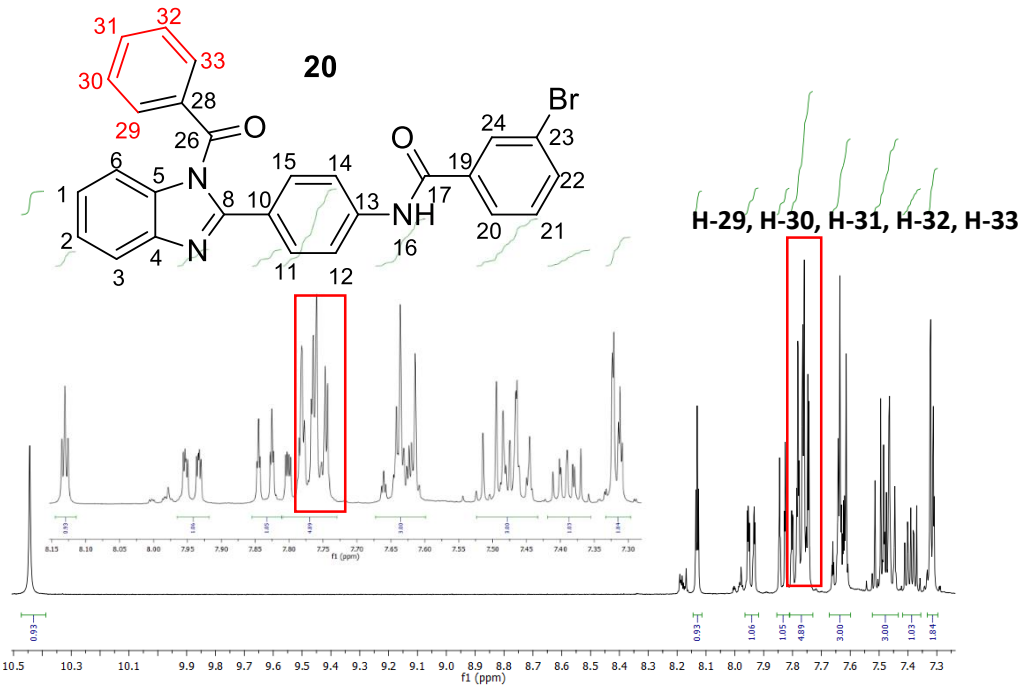


Figure 3.9: The ^{1}H NMR spectrum of **20** in DMSO-d_6, with the E-ring proton signals shown in red.

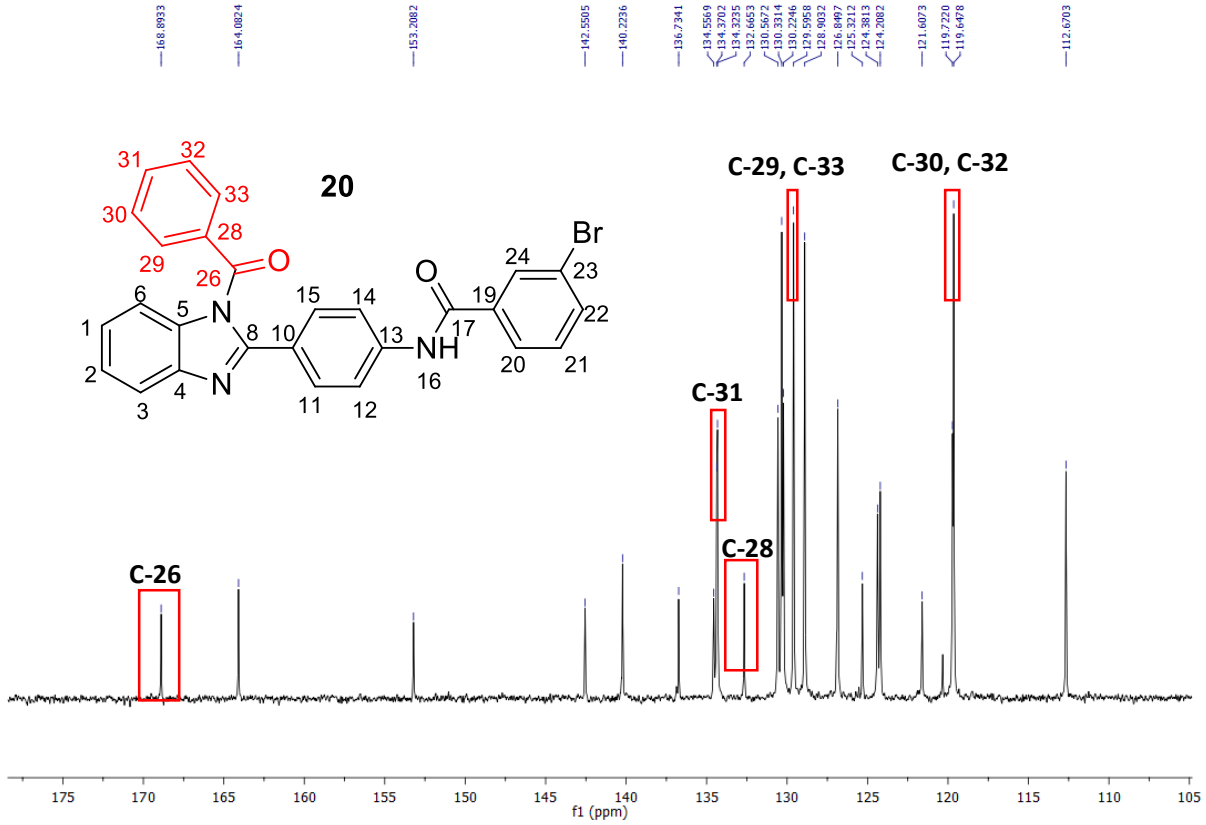
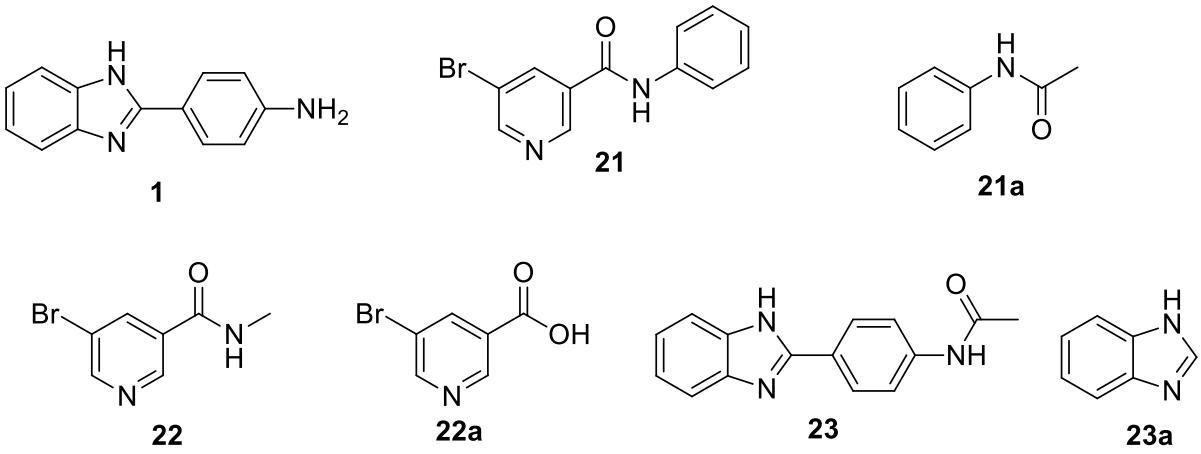


Figure 3.10: The ^{13}C NMR spectrum of **20** in DMSO-d_6, with the E-ring carbon signals shown in red.

3.5.1.4 Benzimidazole Fragment Synthesis

An important issue in any SAR investigation is to evaluate the biological activity of hit fragment structures so as to ascertain at what structural point activity gets switched on. Although many such fragments were predicted by the Bayesian model to be inactive against malaria parasite growth inhibition yet active against β-haematin inhibition, it was deemed necessary to synthesize these compounds in order to verify these results. As with the previous syntheses, the methodologies were well known and are described herein by discussing each compound in turn. Scheme 3.23 shows the seven fragment targets that were purchased (**21a, 22a** and **23a**) or synthesized (**1**, **21**, **22** and **23**) in this study, in which the synthesis of **1** has already been described.

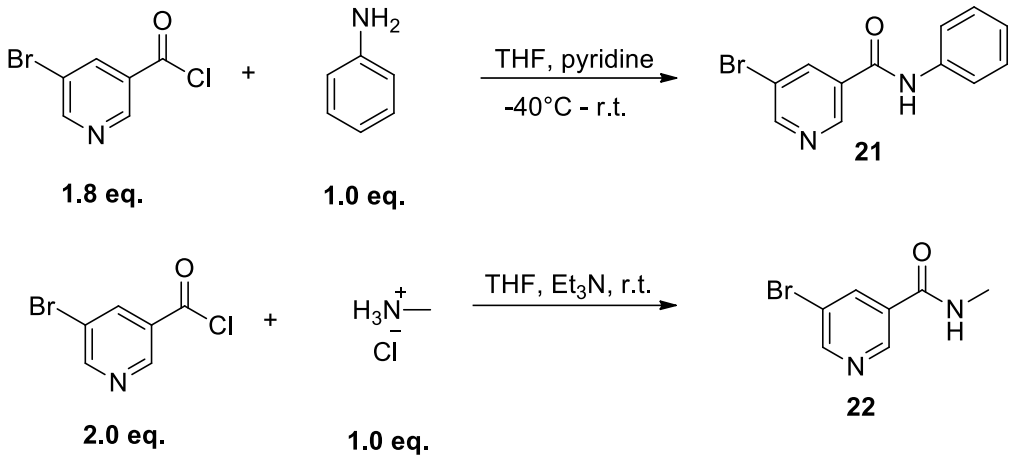


Scheme 3.23: The fragment compounds either purchased or synthesized.

Compound **23** was formed via a standard acylation of **1** by refluxing **1** (1.0 eq.) in acetic anhydride (1.1 eq.) and toluene for two hours at 115 °C. Chemoselective mono-acylation was achieved by controlling the number of reagent equivalents as well as not adding pyridine as a base and acyl-transfer catalyst. The latter resulted in a much slower reaction that required heat to proceed and in which the benzimidazole was presumably just too non-nucleophilic for acylation to occur.

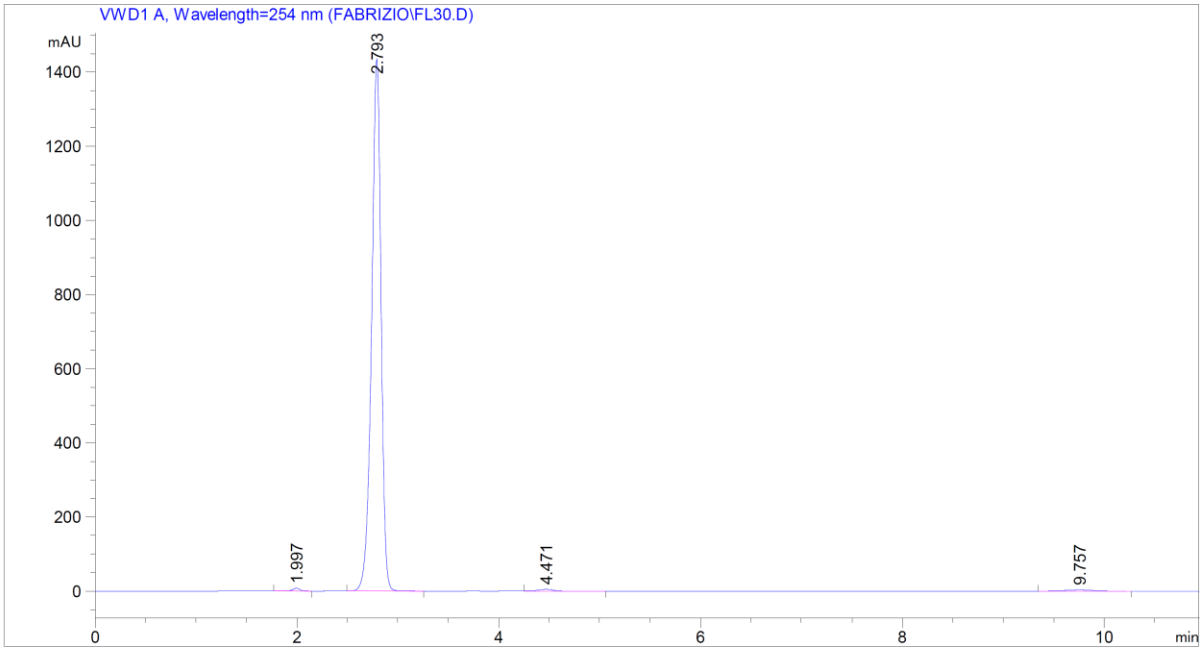
In this case, it was possible to purify the product by simply washing the resulting solid with hot ethyl acetate and toluene to remove excess anhydride and acetic acid by-product. Its ^{1}H NMR spectrum showed exactly the same peaks as **1**, except for the new acetyl group methyl singlet at 2.09 ppm. Similarly, the ^{13}C NMR spectrum showed a new peak at 24.1 ppm accounting for the methyl carbon and a peak at 168.6 ppm resonating for the amide carbonyl. In this case, three A-ring resonances for the benzimidazole carbons in the ^{13}C NMR spectrum demonstrated a symmetrical ring system due to fast H-tautomerism.

Similarly, compounds **21** and **22** were both synthesized using the acid chloride methodology already discussed, but with no risk for formation of a diacylation product. The reaction conditions for both reactions are shown in Scheme 3.24.



Scheme 3.24: Reaction conditions for the formation of compounds **21** and **22**.

For the formation of compound **22**, the methylamine hydrochloride salt was neutralized using triethylamine (Et_3N) to liberate the free amine. After reaction completion, both compounds **21** and **22** were found to be non-polar enough for standard extraction procedures using either chloroform / MeOH mixtures (**22**) or DCM (**21**) after acidification using 1M HCl to pH 7. This was followed by column chromatography and recrystallization using DCM (**21**) or acetone (**22**) to yield the pure products, with purities of greater than 95% according to HPLC. The ^{1}H NMR spectrum of **21** showed six separate resonances accounting for the two aromatic rings systems, as well as an amide peak at 10.47 ppm as expected. The ^{13}C NMR spectrum showed ten expected carbon signals. Similarly, the ^{1}H NMR spectrum of compound **22** showed three proton resonances accounting for the pyridine ring system, a peak at 7.96 ppm accounting for the secondary amide proton and a methyl doublet at 2.93 ppm. Its ^{13}C NMR spectrum showed the expected seven carbon peaks. As with the initial benzimidazole series, all compounds were also characterized using IR spectroscopy, melting point analysis and C18 HPLC purity analysis (Figure 3.11).




```
=====================================================================
                        Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier:                   :      1.0000
Dilution:                     :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=254 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   1.997 VB    0.0777   38.50311    7.39884   0.4044
   2   2.793 BB    0.0993 9381.85059 1433.78235  98.5425
   3   4.471 VB    0.1543   44.42367    4.41759   0.4666
   4   9.757 BB    0.2827   55.83930    3.07723   0.5865
```

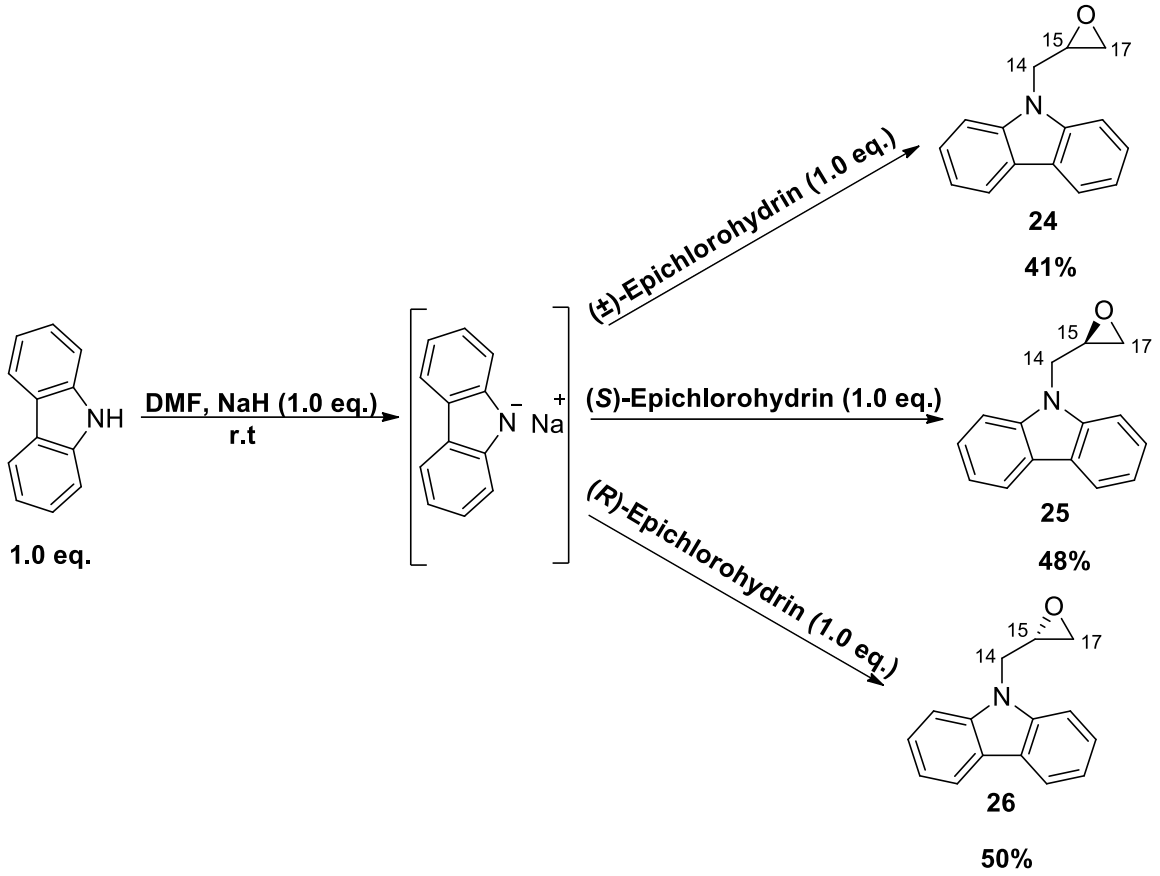
Figure 3.11: C18 HPLC trace of compound **21** at 254 nm.

3.5.2 Scaffold 2: Carbazole Indoles

3.5.2.1 Synthesis of Carbazole and Indole Intermediates

To reach the goal of synthesizing the seven carbazole indole compounds, the synthesis of the four key intermediates (**24**, **25**, **26** and **27**) needed to be realized. The three epoxide intermediates (**24**, **25** and **26**), shown in Scheme 3.25, were synthesized using a previously established methodology reported by Kimura *et al*[157], via alkylation of the respective epichlorohydrin stereoisomer with the carbazole anion derived from using sodium hydride as the base, the latter resulting in a colour change from light yellow to dark purple. Following

the addition of the epichlorohydrin, the reaction proceeded over sixteen hours to allow for complete conversion. The products were purified using column chromatography and recrystallized in ethanol to form long needle-like crystals.



Scheme 3.25: Reaction conditions for the formation of compounds **24**, **25**, and **26**.

As expected, the ^{1}H NMR (Figure 3.12) and ^{13}C NMR spectra of compounds **24**, **25** and **26** were found to be identical, revealing four resonances due to symmetry of the two rings accounting for all the carbazole aromatic protons, two double doublets accounting for the enantiotopic protons H-17, while two double doublets accounted for the enantiotopic protons H-14 and a characteristic epoxide multiplet accounted for H-15. The ^{13}C NMR spectra revealed nine carbon singlets, six for the two identical phenyl ring systems and three for the alkyl chain, confirming successful synthesis of the epoxide derivatives.

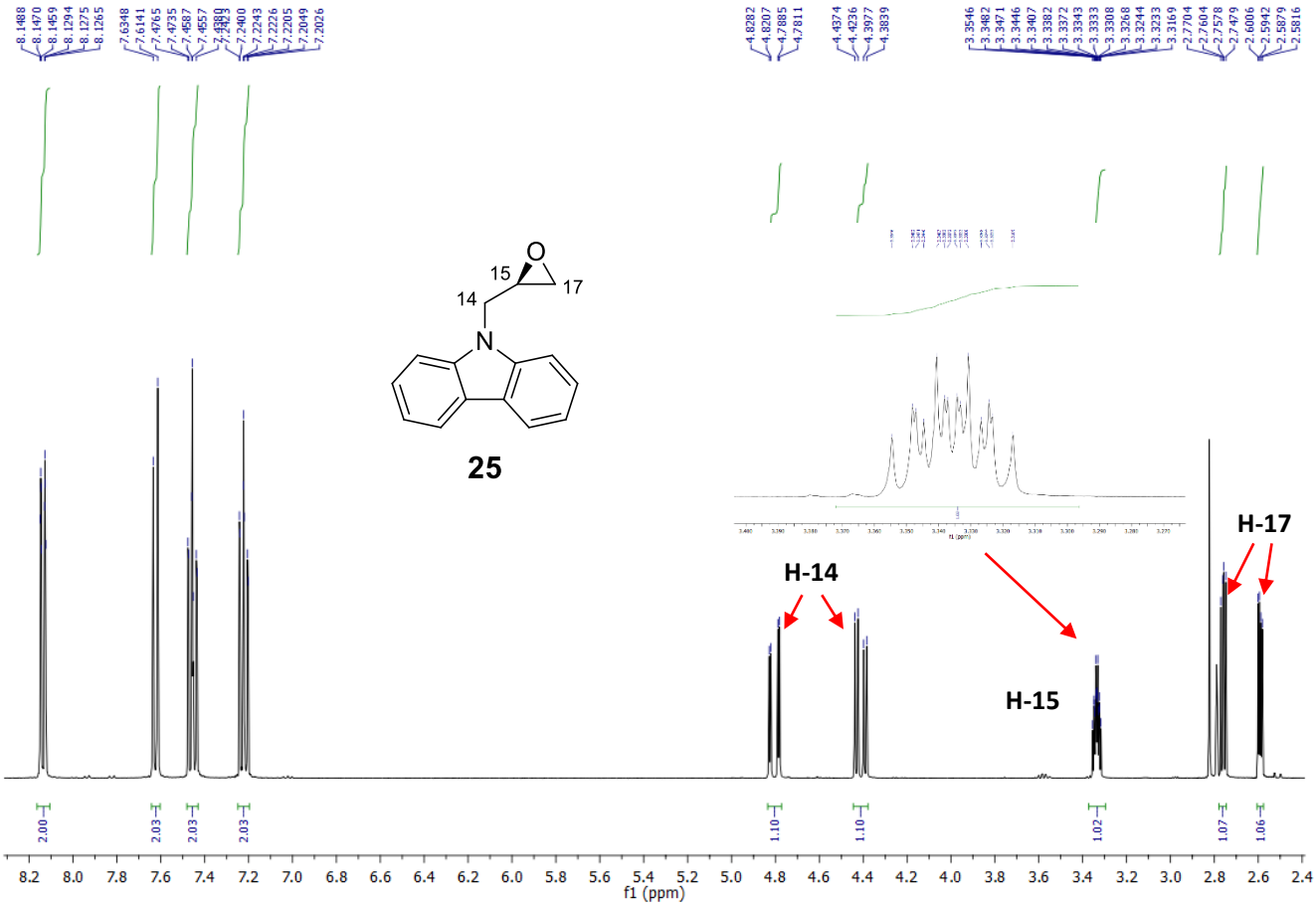
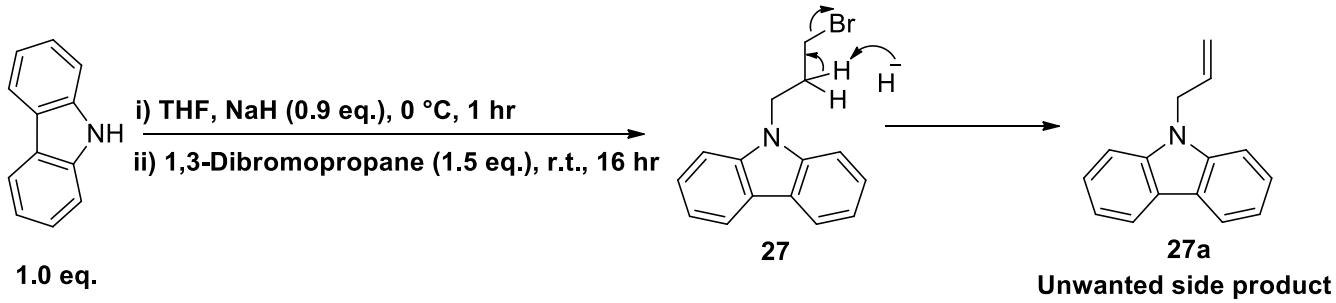


Figure 3.12: The ^{1}H NMR spectrum of **25 in** acetone-d_6 with red arrows indicating the enantiotopic protons and the zoomed in area of H-15.

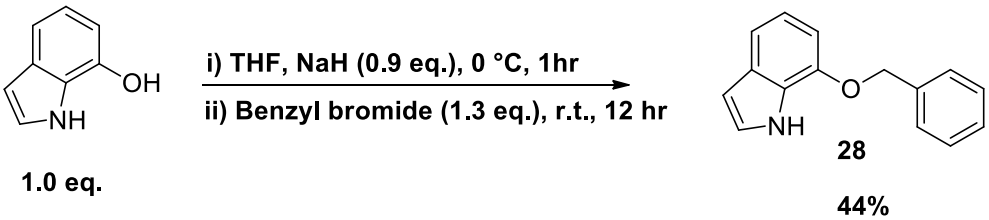
A similar S_N2 alkylation approach was used to synthesize compound **27** as shown in Scheme 3.26, using 1,3-dibromopropane (1.2 eq.) as the electrophile. Again, the carbazole anion was used, although this time NaH was used as the limiting reagent (0.9 eq.) to avoid elimination of the terminal bromide post-alkylation. Unfortunately, the reaction still furnished a 50:50 mixture of the desired bromide **27,** as well as its elimination product **27a**, which was confirmed by ^{1}H NMR spectroscopy. No double-alkylation product was observed though. Compound **27a** was observed presumably due to residual base in the reaction mixture, implying that the NaH had not been fully consumed in spite of it being the limiting reagent as well as, what was considered to be adequate time for deprotonation of 30 minutes at 0 °C followed by 30 minutes at ambient temperature. This was a minor setback, though as both carbazole and 1,3-dibromopropane are cheap and readily available and thus it was not deemed necessary to optimise this reaction. Chromatographic separation of **27** and **27a** was achieved using solely hexane. The ^{1}H NMR spectrum of **27** was very similar to that of **24**, **25**

and **26** as it showed four aromatic proton resonances for the carbazole ring system. Three unique resonances as two triplets and a pentet accounted for the alkyl chain of **27**.



Scheme 3.26: The reaction conditions for the formation of **27** and the mechanism for the formation of **27a**.

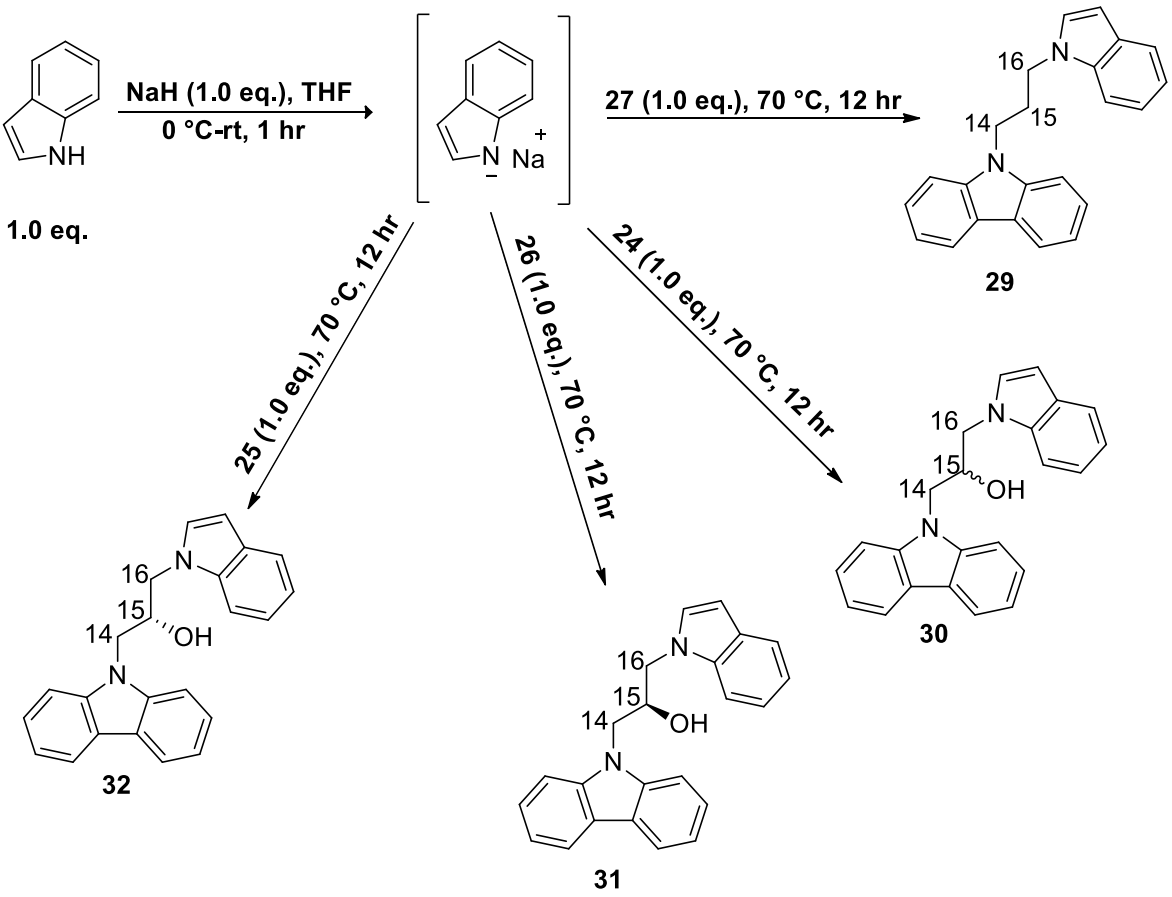
The preparation of compound **28**, as shown in Scheme 3.27, was also an essential requirement of achieving the three final compounds. Compound **28** was prepared using the same S_N2 mechanistic approach as the other four intermediate compounds described above. Less than one equivalent of sodium hydride was used, which ensured formation of a mono-benzylated ether product using benzyl bromide as an electrophile. This was done since the phenolic hydroxyl group of 7-hydroxyindole is known to be more acidic than its secondary amine. The ^{1}H NMR spectrum of **28** showed six aromatic peaks as five for the indole ring system, and a multiplet integrated for five protons which accounted for the phenyl ring. The presence of a peak at 5.12 ppm, integrated for two protons, confirmed the presence of a benzylic methylene group. The absence of a hydroxyl peak at 4.54 ppm and the presence of the secondary amine peak at 8.38 ppm indicated successful chemoselective protection on the hydroxyl group only.



Scheme 3.27: Reaction conditions for the formation of **28.**

3.5.2.2 Synthesis of Carbazole Indole Compounds

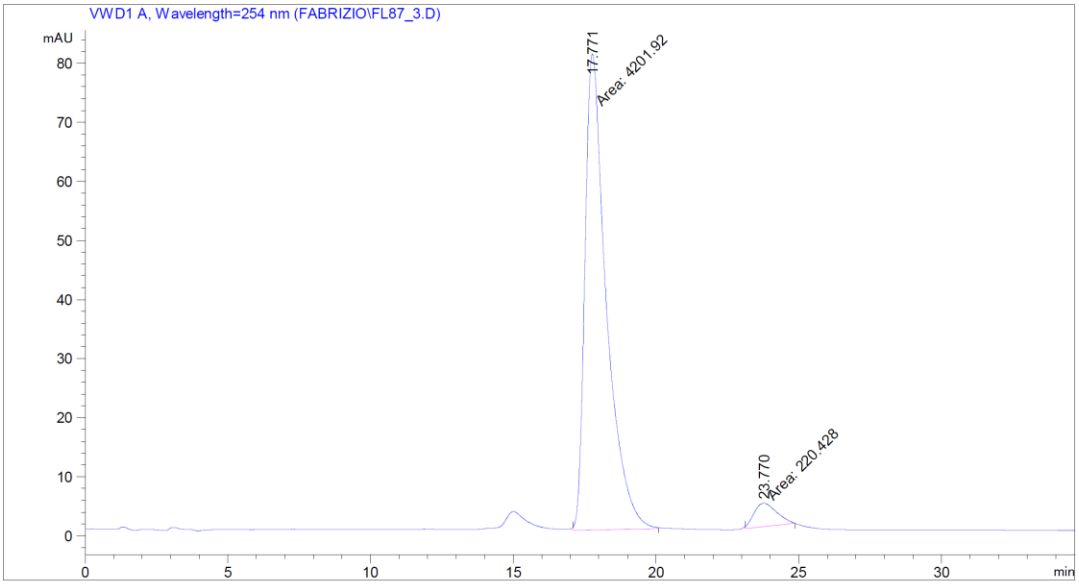
With intermediates **24**-**27** discussed above now in hand, synthesis of the target compounds was carried out involving substitution by an indole anion. The same reaction procedure for synthesis of **24**-**27** except at higher temperature (70 °C rather than the initial room temperature) was used for the formation of compounds **29**, **30**, **31** and **32** as shown in Scheme 3.28. The higher temperature presumably reflected the reduced electrophilicity of the epoxide (Cl replaced by carbazole) as well as the steric presence of the carbazole.



Scheme 3.28: Reaction conditions for the formation of compounds **29**, **30**, **31** and **32.**

As with all NaH reactions, the reaction mixture was quenched in each case with a saturated NH_4Cl solution, and extracted using EtOAc. The compounds **29**, **30**, **31** and **32** were each purified using column chromatography, but an unknown yellow oil was present in each crude sample after column chromatography that could not be separated, as both compounds had the exact same Rf values. The problem was solved by taking advantage of the oily nature of this impurity as well as the crystalline nature of the products by washing the residues with

pentane followed by recrystallization with pentane. This yielded final compounds **29**, **30**, **31** and **32** with moderate yields (51-72%) and, importantly, purities ranging between 93.0 and 99.7% according to HPLC using a C18 column was observed. Compounds **31** and **32** were further subjected to chiral OD HPLC, with enantiomeric excess (ee) of each compound calculated using the racemic mixture (**30**) as a reference. The ee of **31** and **32** were found to be 90 and 80% respectively, with the chiral OD HPLC trace for **31** shown in Figure 3.13. The peak at 15 min was rationalized as a solvent peak. Even though these ee values are not ideal, they were expected to give a strong indication of which enantiomer is the most biologically active after further testing.




```
======================================================================
                        Area Percent Report
======================================================================

Sorted By             :      Signal
Multiplier:                   :      1.0000
Dilution:                     :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=254 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1  17.771 MM    0.8685 4201.92041   80.63261  95.0156
   2  23.770 MM    0.9230  220.42789    3.98041   4.9844
```

Figure 3.13: The chiral OD HPLC trace of compound **31**.

As expected, the ^{1}H NMR spectra of the racemic mixture (Figure 3.14) and enantiomers (**30**, **31**, **32**) were identical, with ten resonances accounting for the both the carbazole and indole

ring systems. However, proton signals were overlaid onto each other, making it difficult to properly distinguish between the aromatic signals. Additionally, diagnostic peaks from the C-3 spacer included a broad singlet at 1.99 for the hydroxyl group, two multiplets, integrated for two protons each accounting for the protons H-14 and H-16, and a multiplet integrated for one proton accounting for H-15.

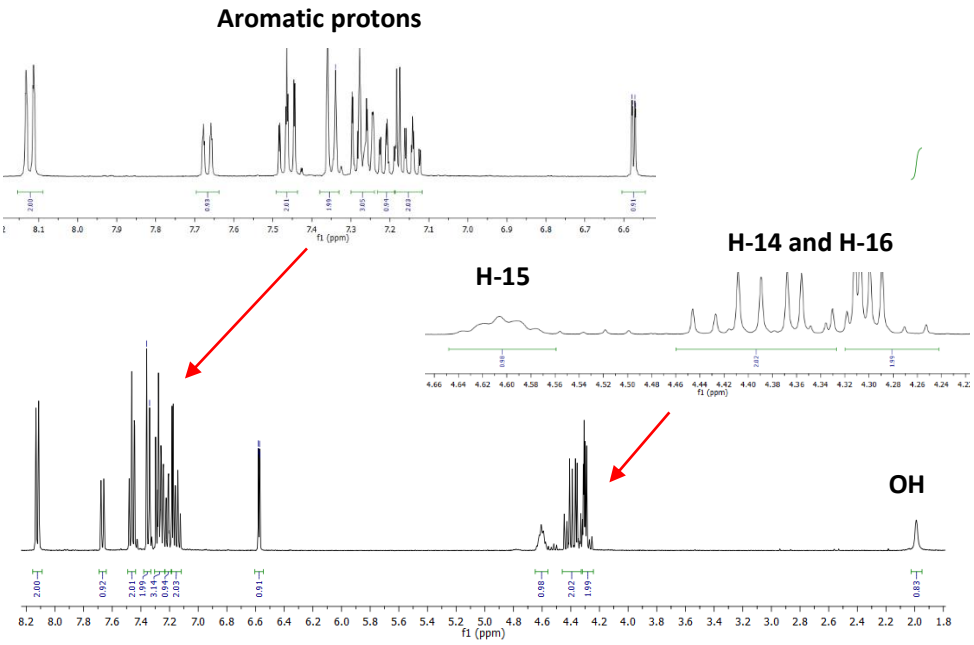


Figure 3.14: ^{1}H NMR spectrum of **30 in** chloroform-*d* with the red arrows indicating zoomed in areas.

The ^{13}C NMR spectrum showed seventeen carbon peaks containing the three diagnostic alkyl peaks at 47.2, 50.4 and 70.6 ppm respectively, and the remaining fourteen peaks which accounted for the aromatic carbons.

The aromatic region of the ^{1}H NMR spectrum of **29** revealed a similar pattern as for **30**-**32**, but the absence of the stereogenic hydroxyl-bearing centre made the spectrum much simpler in the aliphatic region, with just two triplets and one pentet observed. The ^{13}C NMR spectrum of **29** was also found to be similar, with seventeen carbon peaks, in which the alkyl carbons were shifted upfield at 29.4, 40.5 and 44.0 ppm respectively due to the lack of the electron-withdrawing secondary alcohol group. All compounds were also characterized using IR spectroscopy, melting point analysis, HPLC and HRMS, with the HRMS analysis for compound **29** shown in Figure 3.15 as an $[M+H]^{+}$ molecular ion using electrospray positive ionization.

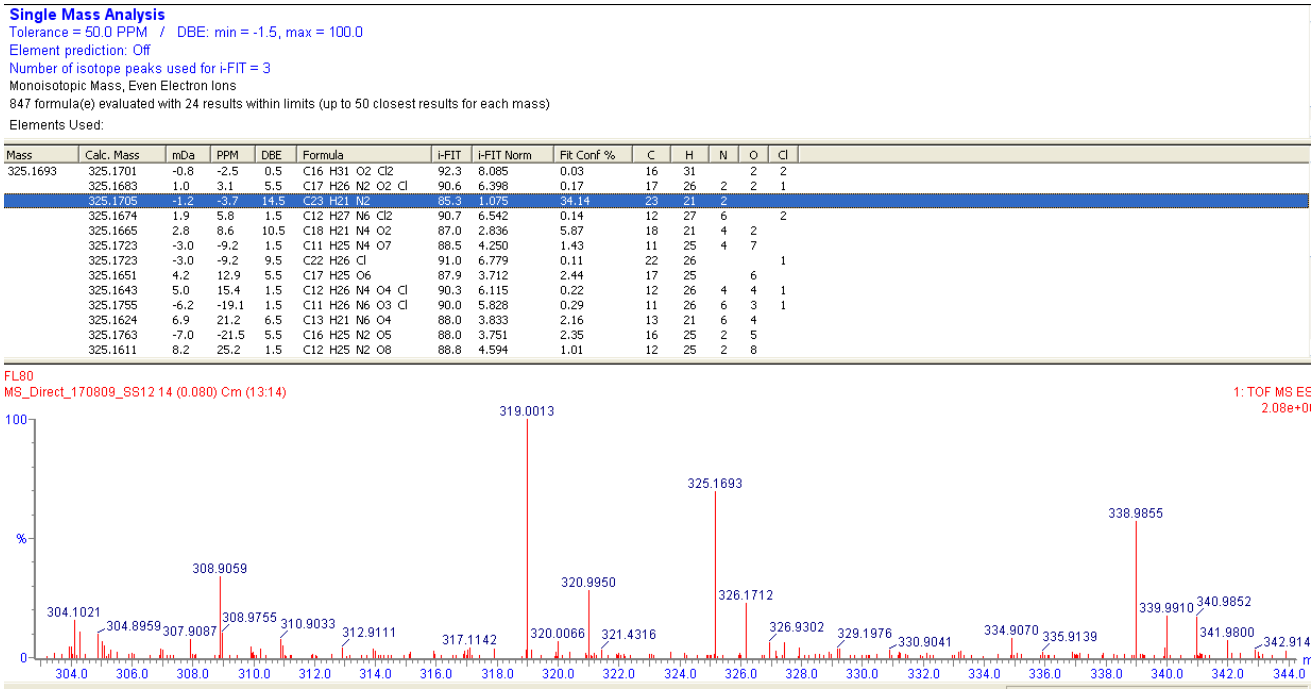
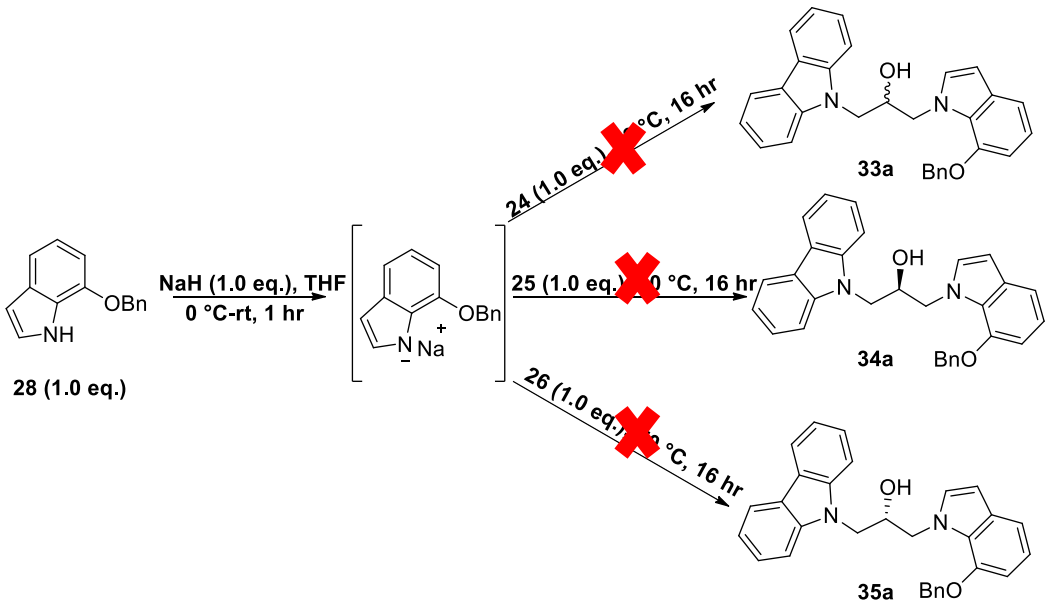


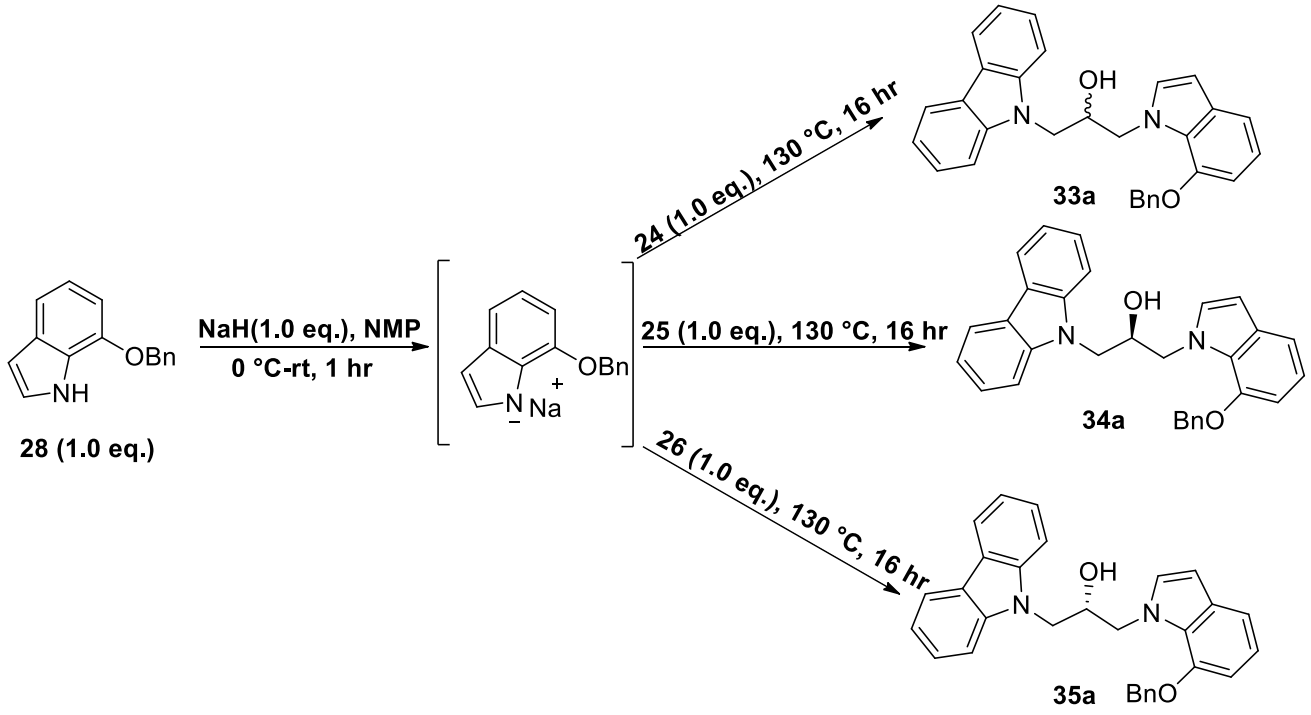
Figure 3.15: The HRMS analysis of compound **29** using electrospray positive ionization.

Since the formation of the intermediate compounds **33a**, **34a** and **35a** from benzyl ether **28** followed the same mechanistic approach as that of formation of compounds **29**, **30**, **31** and **32,** it was decided to use the same methodology (Scheme 3.29). Although, at 70 °C the reaction showed no formation of product which was attributed to the large benzyl protecting group of compound **28,** providing steric hindrance**.**



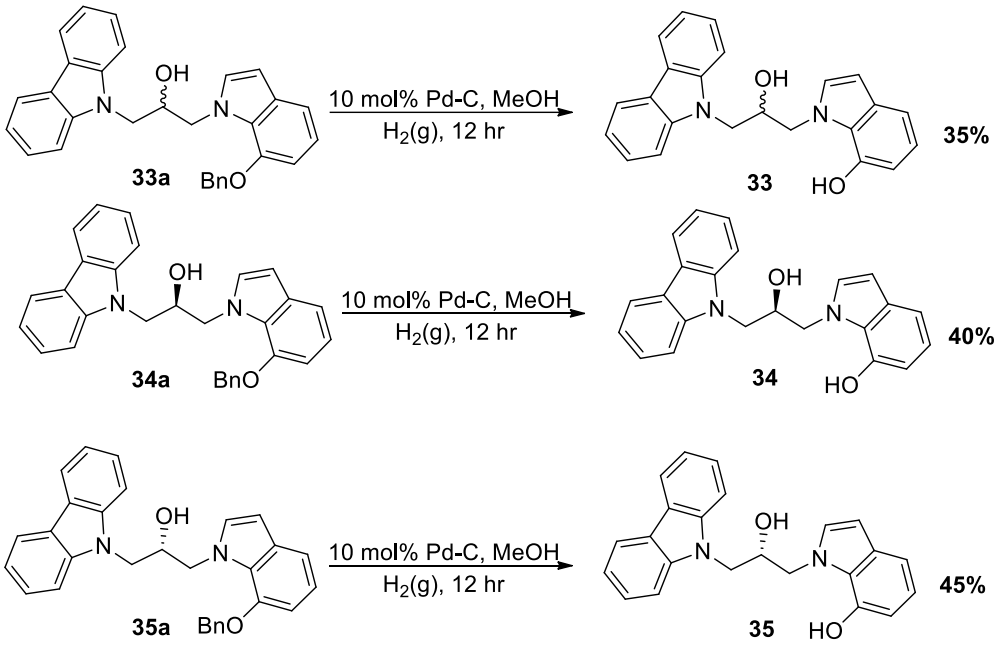
Scheme 3.29: Reaction conditions for the first attempt at formation of compounds **33a**, **34a** and **35a** using the previously established methodology.

This was addressed by increasing the temperature to 130 °C and changing to DMF as a solvent. However, this choice also proved to be problematic as the starting material degraded at high temperature under basic conditions according to TLC. Hence, *N*-methyl-2-pyrrolidone (NMP) was used as a replacement solvent with some success but this too had limitations in that the reaction failed to achieve full conversion. After further optimization attempts using NMP, including increasing the temperature from 130 to 200 °C, the number of NaH equivalents from one to two and using longer reaction times (up to one week) there was still no improvement in the yield. Ultimately, the final reaction conditions settled on are shown in Scheme 3.30. It was decided not to purify these compounds to the usual standards (≥ 95% purity at this stage, but rather deprotect the benzyl groups first before fully characterizing. Therefore, a minimum purification (removal of excess starting materials and NMP) was carried out in each case whereby, the reaction mixture was quenched with a saturated NH_4Cl solution and extracted using EtOAc. The organic residue was washed nine times with water to remove all the NMP and purified using column chromatography.



Scheme 3.30: The reaction conditions used for the synthesis of intermediates **33a, 34a** and **35a**.

Following this, the benzyl protecting groups of **33a, 34a** and **35a** were each cleaved hydrogenolytically using 10 mol% of Pd-C and hydrogen gas in anhydrous methanol to furnish the final compounds **33**, **34** and **35** as shown in Scheme 3.31. The reaction mixtures were each filtered through Celite©, taking care not to let the Pd-C residue run dry for fear of ignition. The residues were purified using column chromatography, and as with the previous compounds **29**-**32**, the yellow oily impurity was removed by washing the residues with pentane. The remaining solids were recrystallized using DCM to yield compounds **33**, **34** and **35,** with chemical purities of greater than 95% according to HPLC on a C18 column. The yields (35-45% over two steps) for this subset of compounds were relatively poor, due to incomplete conversion and difficult purification.



Scheme 3.31: Reaction conditions for the benzyl deprotection reactions to furnish the final compounds **33, 34** and **35** with yields given for over two steps.

The ^{1}H and ^{13}C NMR spectra of compounds **33-35** were identical and needed to be carried out in methanol-d_4 due to low solubility in other solvents. Since the visualization of the hydroxyl groups was not possible using ^{1}H NMR in this solvent, IR spectroscopy was used instead, in which two separate hydroxyl peaks (OH phenolic and OH aliphatic) indicated deprotection and therefore formation of the desired compounds as revealed in Figure 3.16 for compound **33**.

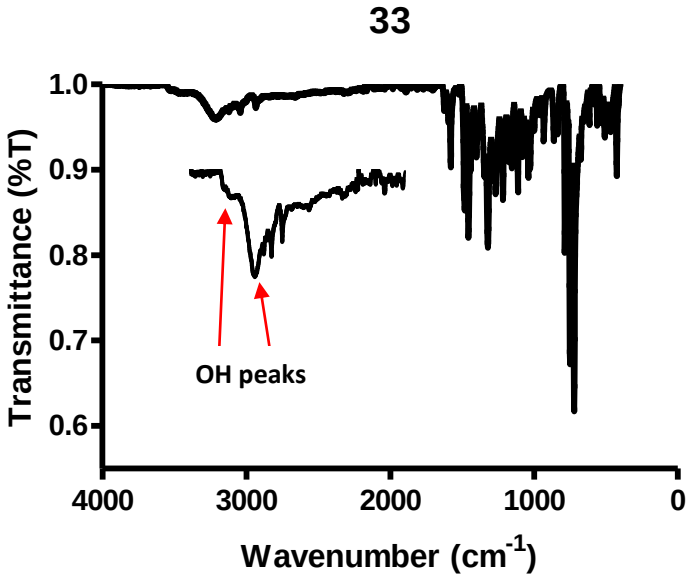


Figure 3.16: The ATR-FTIR spectrum of **33** with the two hydroxyl peaks indicated by the red arrows.

The ^{1}H NMR spectra of **33**-**35** each showed nine separate resonances accounting for the aromatic protons, and five resonances which corresponded to the four diastereotopic protons H-15 and H-17 together with H-16 for the stereogenic centre as shown in Figure 3.17. The ^{13}C NMR spectra of **33**-**35** showed the expected 17 singlet resonances accounting for all carbons present in the molecules. As with the other subsets of compounds, **33**-**35** were also characterised using IR spectroscopy, melting point analysis and HRMS, with an HRMS analysis for compound **35** using electrospray positive ionization shown in Figure 3.18.

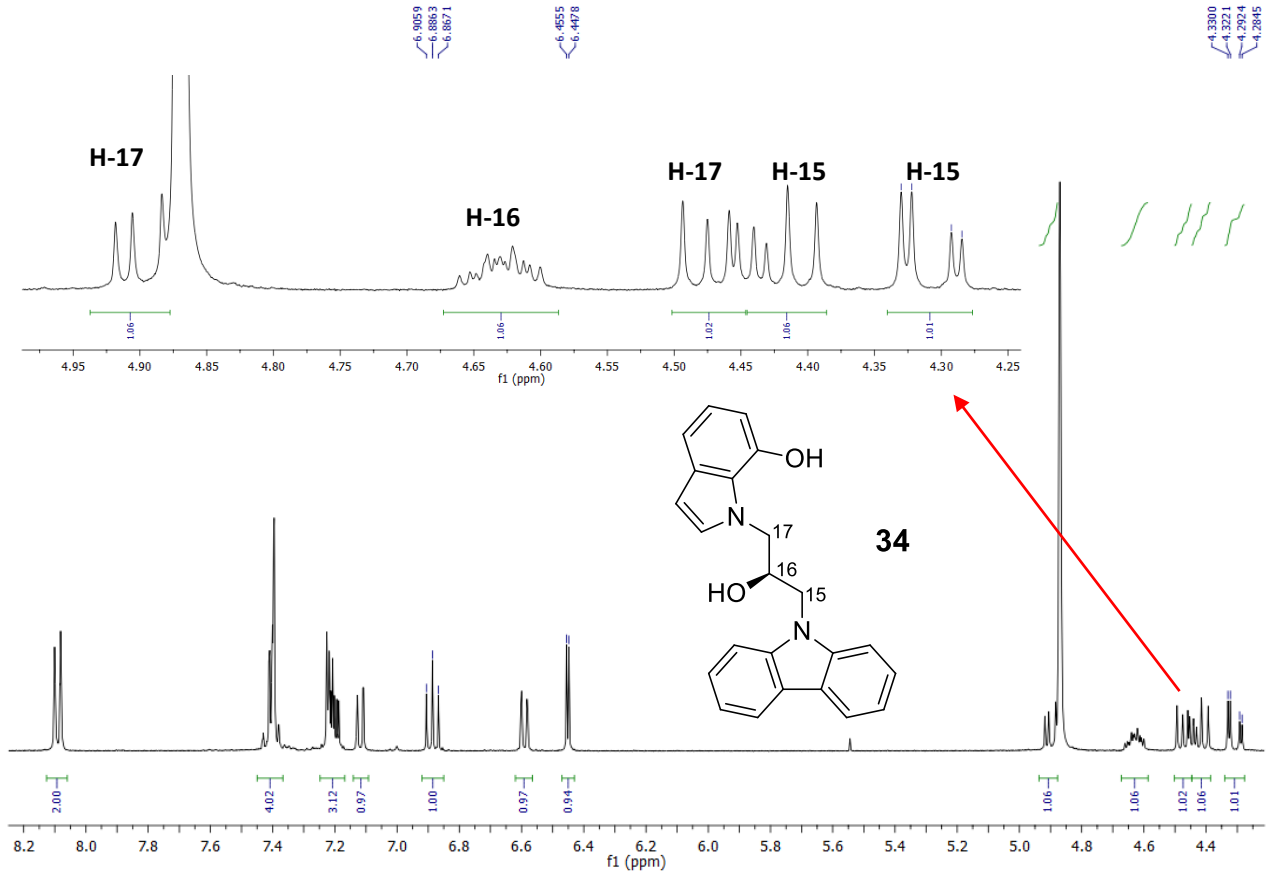


Figure 3.17: The ^{1}H NMR spectrum of **34** in methanol-d_4 with the zoomed in alkyl region shown in red.

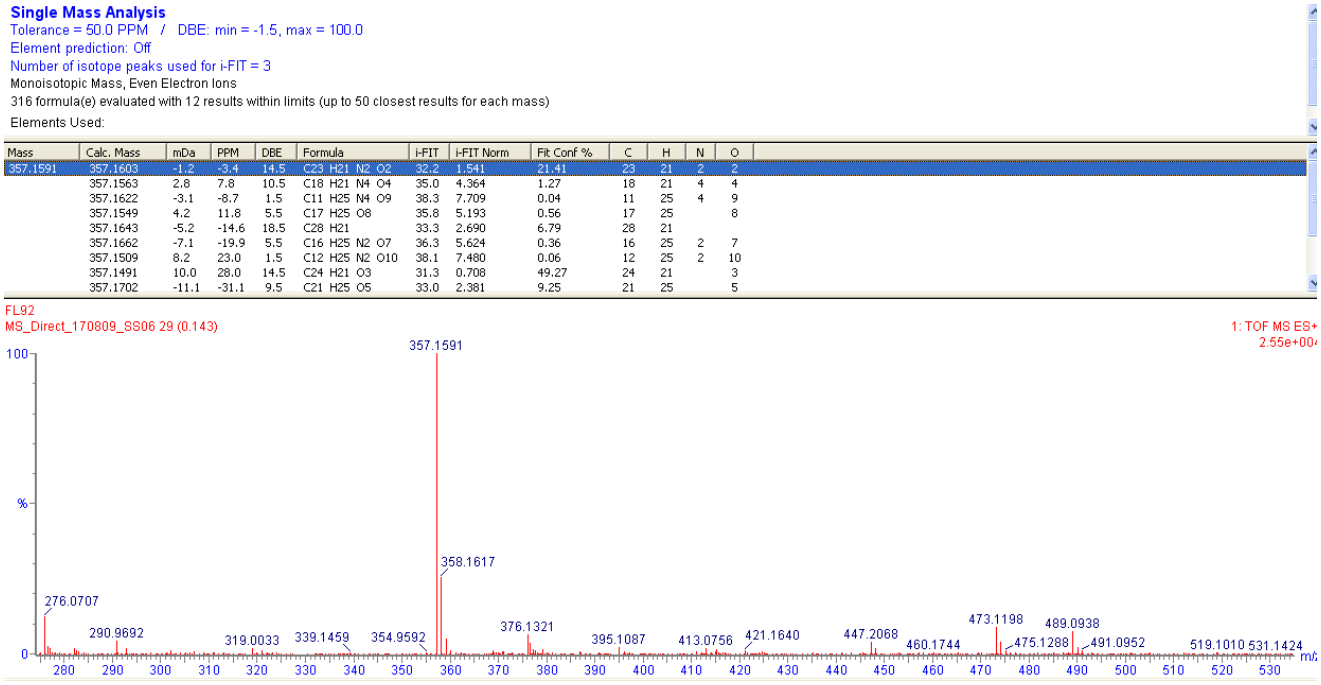
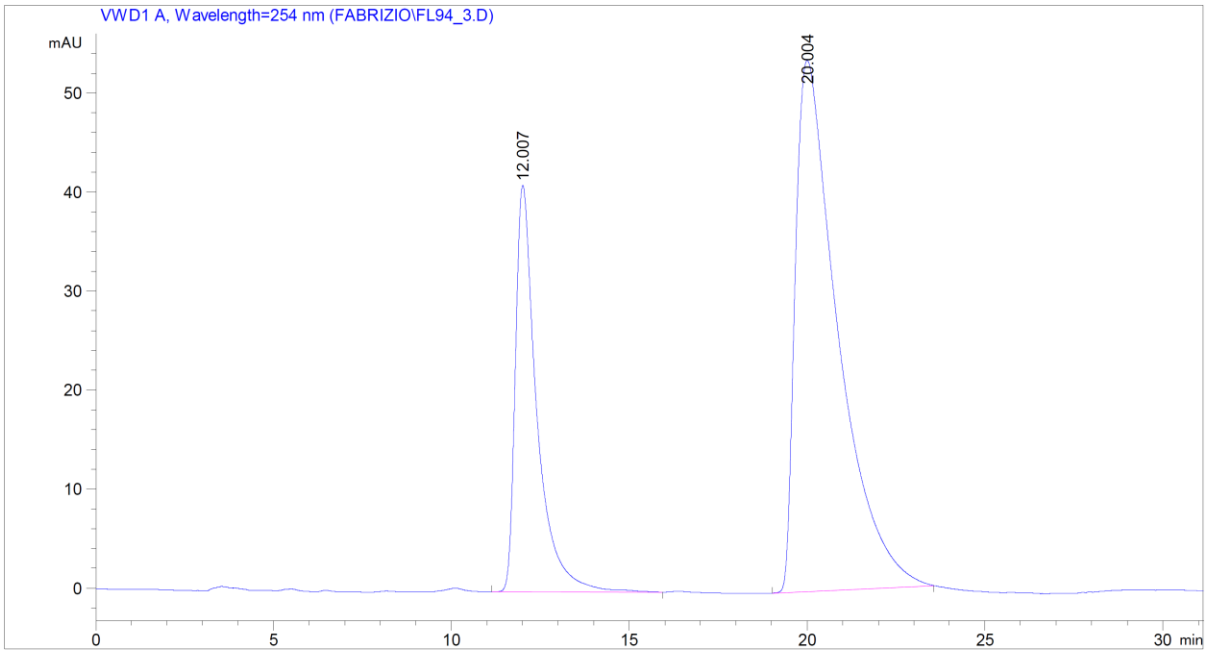


Figure 3.18: The HRMS analysis of compound **35** using electrospray positive ionization.

Compounds **35** and **34** were further subjected to chiral OD HPLC with the ee of each compound calculated using the racemic mixture (**33**) as a reference. The ee of **34** and **35** were found to be 80% and 44% respectively with the HPLC trace of **35** shown in Figure 3.19. The stereoselectivity of **34** and **35** was found to be worse than that of **31** and **32,** presumably due to the increased temperature required for reaction completion, although a precise mechanism for racemisation was not clear. These ee values were far from perfect, especially that of compound **35**; however, they were considered to offer an insight for preliminary biological studies where the objective was to identify the most active enantiomer. C18 HPLC analysis of compounds **33**-**35** also showed chemical purities of ≥ 95%.




```
=====================================================================
                      Area Percent Report
=====================================================================

Sorted By             :      Signal
Multiplier:                   :      1.0000
Dilution:                     :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=254 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1  12.007 BB    0.6127 1727.92627   41.08024  27.9896
   2  20.004 BB    1.1287 4445.53076   53.64973  72.0104
```

Figure 3.19: The chiral OD HPLC trace of compound **35**.

3.6 Summary and Conclusions

A synthetic template, **1**, was synthesized using PPA, 4-aminobenzoic acid and *o*-phenylenediamine using known literature methodology for intended elaboration into a library of benzimidazole-tethered amides. All attempts at formation of A-ring derivatives of **1** using the same methodology failed due to degradation, in spite of temperature optimization studies. By comparison, the amino group of **1** could be used to synthesize a library of compounds **2**-**19** via acid chloride methodology with pyridine as a base and an acyl-transfer agent, with THF as the solvent. This process was optimised using two equivalents of acid chloride at a temperature of -40 °C to avoid formation of any diacylated product. A diacylated product, **20**, was also synthesized to compare biological activities with the parent compound, **3**. The benzimidazole fragment compounds were either purchased (**21a**, **22a**, **23a**), preformed (**1**) or synthesized using the acid chloride methodology (**21**, **22**) or standard acetylation methodologies (**23**).

Formation of carbazole epoxide intermediates (**24**, **25** and **26**) was achieved via a NaH (1.0 eq.) mediated S_N2 reaction using both chiral and racemic epichlorohydrin reagents (1.0 eq.) individually, and carbazole (1.0 eq.), with moderate yields of 41 to 50%. Formation of intermediate **27** involved changing the electrophile to 1,3-dibromopropane, with a reduction in the number of equivalents of NaH used in the hope of limiting formation of **27a** via elimination of bromide. However, this was to no avail, implying unreacted NaH was still present in the reaction mixture upon addition of the 1,3-dibromopropane. These intermediates (**24**, **25**, **26**, **27**) were subjected to a further NaH (1.0 eq.) mediated S_N2 reaction using indole (1.0 eq.) as the nucleophile to furnish compounds **29**, **30**, **31** and **32** with moderate yields of 54 to 72%. A problematic yellow oily residue present in compounds **31**, **32** and **33** made purification difficult, which was removed by washing and recrystallization of the residues with pentane. The enantiomeric excesses for **31** and **32** were found to be 90 and 80% respectively, which was good enough for taking through to biological testing so as to give an indication of enantiomer relative activity.

The benzyl protection of 7-hydroxyindole via a NaH mediated S_N2 reaction furnished intermediate **28.** Using **28** as a reagent, while modifying the NaH reaction conditions by changing the solvent from THF to NMP and increasing the reaction temperature from 70 to

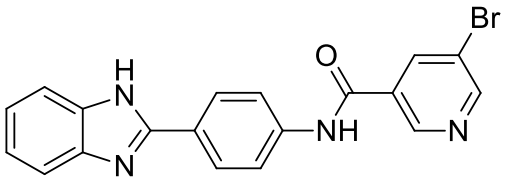
130 °C, successfully produced intermediates **33a**, **34a** and **35a**. Benzyl group hydrogenolysis of **33a**, **34a** using Pd-C, MeOH and in a hydrogen atmosphere furnished compounds **33**, **34** and **35,** with yields of 33 to 45% over two steps**.** The enantiomeric excess of **34** and **35** was found to be 80 and 44% respectively with loss of enantio-purity attributed to the harsh reaction conditions used.

Discounting intermediates, a total of twenty three benzimidazole compounds were synthesized using no more than two steps, with moderate to excellent yields and chemical purities of ≥ 95%. A variety of derivatives incorporating different functional groups was synthesized for SAR and QSAR analysis. Also discounting intermediates, a total of seven carbazole indole compounds were synthesized using either a two or three-step synthesis, with moderate to low yields and chemical purities ≥ 95%. All characterization data confirmed identity of all synthesized compounds and all compounds were soluble enough in DMSO or via sonification for biological testing.

Chapter 4: Benzimidazole Structure-Activity Relationships: β-haematin Inhibition, Parasite Growth Inhibition and Molecular Docking

4.1 Introduction

After synthesizing both series of benzimidazole compounds, the next logical step was to analyse their β-haematin inhibiting and biological activities to verify the Bayesian statistical predictions. The β-haematin inhibition and parasite growth inhibition activity of the hit compound (Figure 4.1), **5**, was already reported in the Vanderbilt high-throughput screen. Therefore, the aim of this section was to replicate that data and determine the activities of all other synthesized compounds in both benzimidazole series.



β-haematin IC_{50}: 14.3 μM

Parasite growth inhibition IC50: D6 strain 348 nM, C235 strain 412 nM

Figure 4.1: The reported activities of the hit benzimidazole compound, **5**.

Structure-activity relationships (SARs) have been used widely in the literature to predict biological or binding data based on physiochemical properties of compounds from a scaffold that contains previous activity data.[138] This allows the user to predict biological activities pre-synthesis, making it a more efficient and rational process, thus saving in terms of cost and time. Molecular docking is also a useful computational tool to predict interactions with molecules and crystal or protein surfaces and can help confirm the synthesized compounds interaction with the target, in this instance the β-haematin crystal surface.[158] All these techniques are discussed and used in this chapter.

4.2 Benzimidazole Structure-Activity Relationships and Molecular Docking Goals

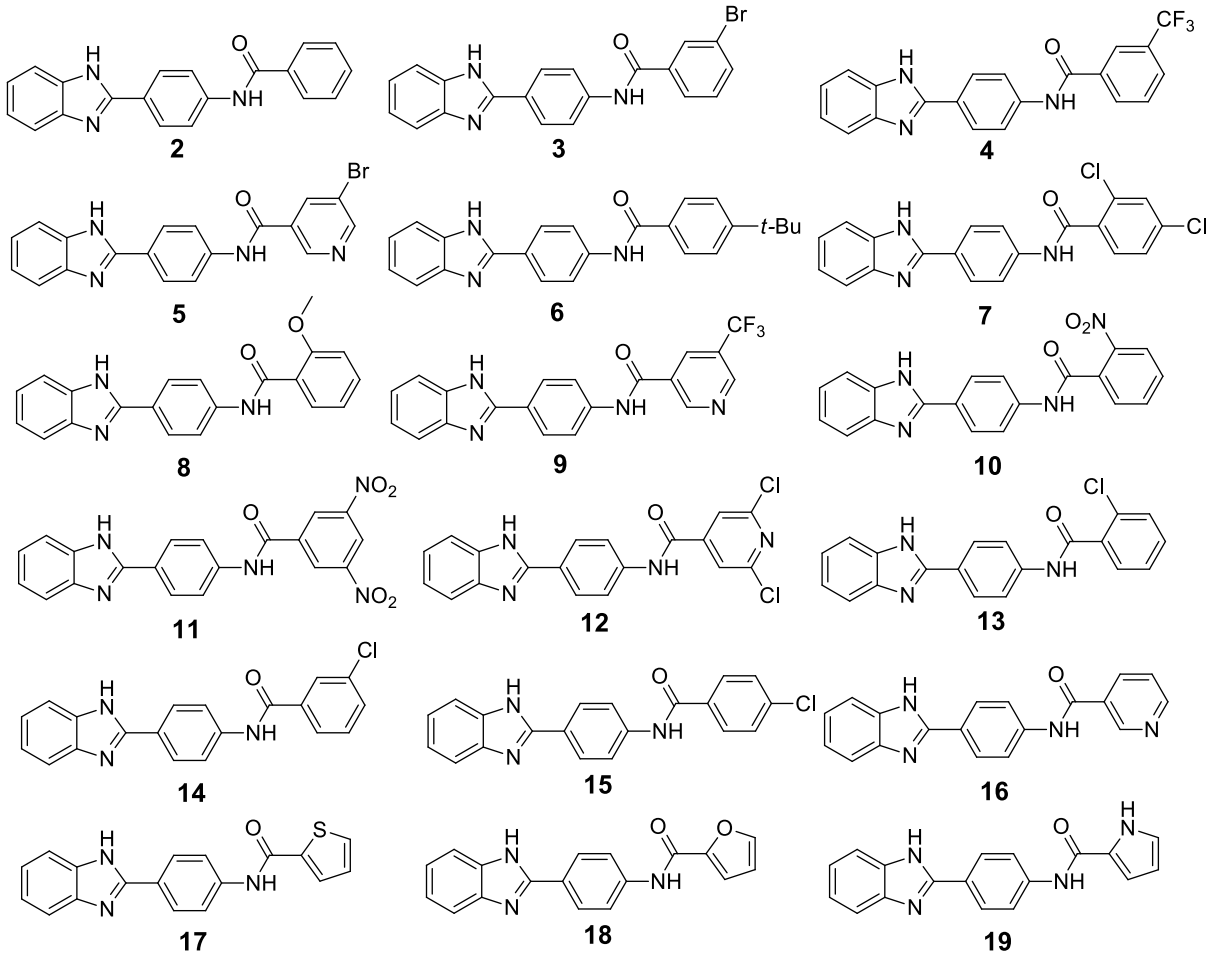
The goals of this chapter were to:

- Perform the NP-40 β-haematin inhibition assay on all benzimidazole compounds.
- Use Materials Studio to perform docking studies on the first subset of benzimidazole compounds (**2**-**19**).
- Subject all benzimidazole compounds to parasite growth inhibition studies on the NF54 strain and analyse the data.
- Subject selected benzimidazole compounds to parasite growth inhibition studies on the Dd2 strain and determine their cytotoxicity.
- Subject selected compounds to the haem fractionation assay.
- Using the data from the above studies, attempt to establish a QSAR for compounds **2**-**19**.

4.3 Results and Discussion

4.3.1 β-haematin Inhibition Studies

The NP-40 detergent mediated assay, modified from Carter *et al*, was used for all these studies. Modifications to improve the solubility of the benzimidazole compounds by substituting 70 µL of water with acetone was necessary to obtain reproducible results. This substitution of solvents has been shown to have no interference in the assay and is in accordance with the published data by Carter *et al.*[92] Compounds **2**-**19** shown in Scheme 4.1 were tested, with the results shown in Table 4.1 accompanied by the Bayesian predictions from Chapter 2.



Scheme 4.1: The synthesized benzimidazole compounds tested for β-haematin inhibition using the NP-40 detergent based assay.

Of the eighteen compounds tested, fourteen were found to be active in the β-haematin inhibition assay using the cut-off of 100 µM for β-haematin inhibition activity with a typical

sigmoidal dose response curve shown in Figure 4.2. This indicated that 78% of the Bayesian predictions were correct. Compound **7** was also found to be very close to the cut-off. Compared to random screening using the β-haematin Bayesian model containing 2 113 active compounds and 64 118 inactive compounds, a 24-fold enrichment was thus achieved.[127]

Table 4.1: The β-haematin inhibition results of compounds **2**-**19** and the accompanying Bayesian prediction from Chapter 2.

Compound	β-haematin inhibition IC_{50} (μM)	Bayesian β-haematin inhibition score
2	Inactive	25.6979
3	24.5 ± 4.2	28.1166
4	28.1 ± 1.7	32.6362
5	32.0 ± 1.4	38.4248
6	Inactive	25.4777
7	Inactive	24.0156
8	38.9 ± 1.5	19.9991
9	16.6 ± 0.3	36.4715
10	28.9 ± 1.6	21.0876
11	13.3 ± 0.5	40.5853
12	16.8 ± 0.4	29.5242
13	38.2 ± 1.4	19.5834
14	25.1 ± 1.1	31.0259
15	Inactive	25.1454
16	35.6 ± 1.7	30.6674
17	42.4 ± 2.1	19.3296
18	84.1 ± 2.4	14.9506
19	46.6 ± 1.7	27.4973

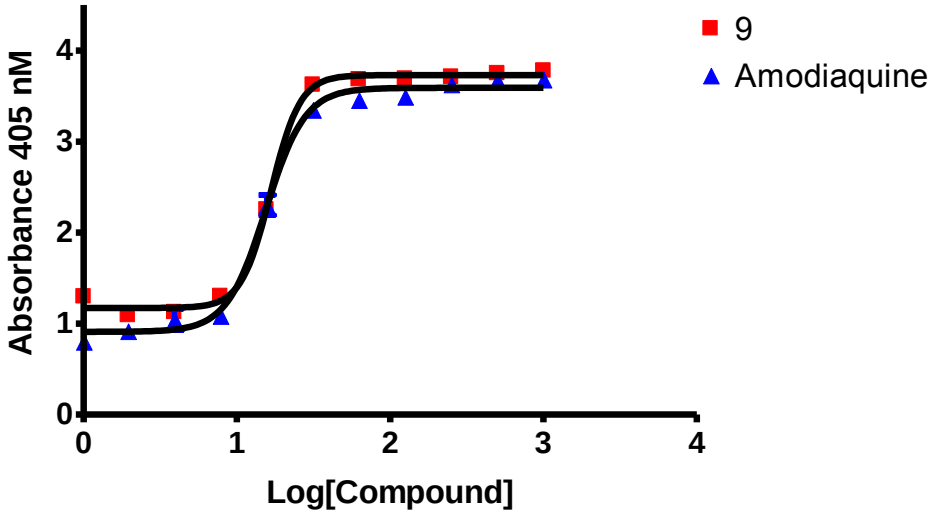
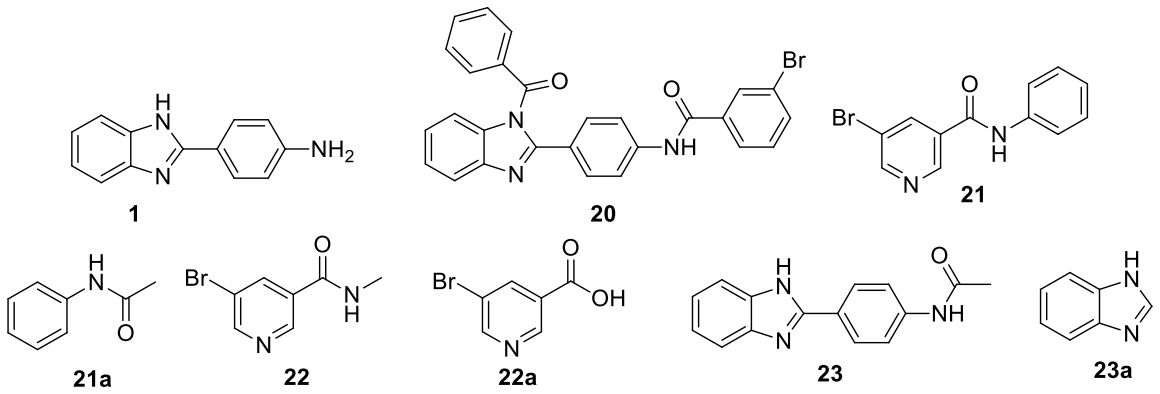


Figure 4.2: A sigmoidal dose response curve showing β-haematin activity of **9** and the amodiaquine standard.

One goal of this study was to determine which fragments of the hit compound, **5**, are essential for β-haematin inhibitory activity. The fragment compounds purchased (**21a, 22a** and **23a**) and synthesized (**1, 21, 22** and **23**) as well as the diacylated product, **20** are shown in Scheme 4.2. These compounds were subjected to the same assay and the results are shown in Table 4.2 along with the Bayesian predictions from Chapter 2. All compounds except **20** were found to be experimentally inactive in contrast to the predictions of six active and one inactive compound. This is not surprising as the Bayesian scores for all compounds except two are very close to the cut-off score for an IC_{50} below 100 µM and hence have low confidence. Compound **20** was found to be experimentally somewhat more active than its parent compound, **3**, with an IC_{50} of 15.6 ± 0.5 µM compared to **3** with an IC_{50} of 24.5 ± 4.2 µM**.** Compound **1** was found to be a false positive on the basis of the β-haematin inhibition activity prediction while compound **23** had a β-haematin inhibition value of 120.6 ± 6.9 µM which is close to the cut-off and accounts for the second highest Bayesian prediction score. Despite all the false positive predictions, this data gave a 14.3% hit rate for these compounds. When this is compared to random screening using the β-haematin inhibition Bayesian model, a 4-fold enrichment was achieved. These two significant enrichments achieved for the benzimidazole compounds (24-fold) and fragment studies (4-fold) show how powerful the Bayesian technique and specifically the β-haematin Bayesian model is as this technique not only saves time but valuable resources when performing hit to lead analysis.



Scheme 4.2: The synthesized benzimidazole fragments.

Table 4.2: The β-haematin inhibition results of the benzimidazole fragment compounds and the accompanying Bayesian prediction, with the above the cut-off score for IC_{50} < 100 µM in red.

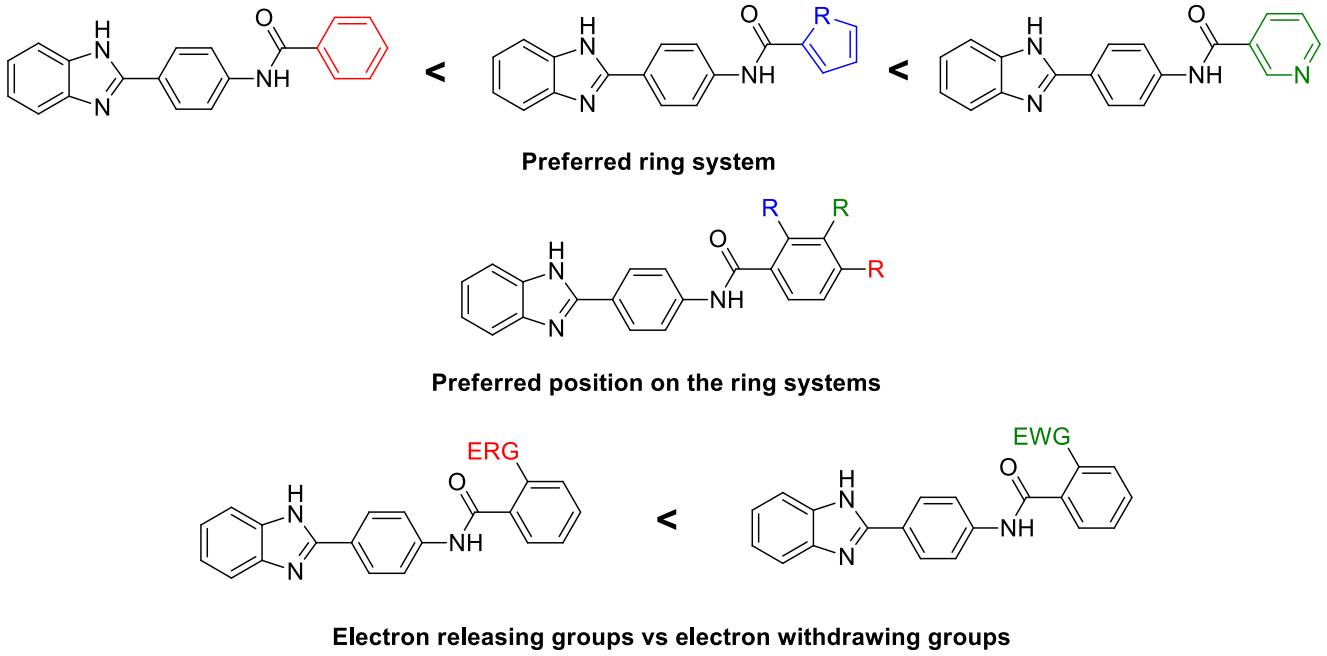
Compound	β-haematin inhibition IC_{50} (µM)	Bayesian β-haematin inhibition score
1	Inactive	22.8883
20	15.6 ± 0.5	-
21	Inactive	12.7431
21a	Inactive	-5.34745
22	Inactive	6.81007
22a	Inactive	8.04823
23	120.6 ± 6.9	19.8929
23a	Inactive	1.03421

4.3.2 Benzimidazole β-haematin Inhibition SAR

When analysing compounds **2**-**19** for SARs, clear trends emerged. By considering the β-haematin activities of compounds **13, 14** and **15,** which are all mono-chlorinated compounds, the effect of substituent position on the phenyl ring system was made evident. Compound **13** (*o*-chloro) showed moderate β-haematin activity of 38.2 ± 1.4 µM, while compound **14** (*m*-chloro) showed good activity (25.1 ± 1.1 µM) and compound **15** (*p*-chloro)

no activity (above 100 μM). Upon examination of the other *para*-substituted compounds **6** and **7**, both compounds showed similar inactivity in the β-haematin inhibition assay. The same holds true for the *ortho*-substituted compounds **8** (*o*-methoxy) and **10** (*o*-nitro), with moderate β-haematin activities of 38.9 ± 1.5 μM and 28.9 ± 1.6 μM. It is evident that the electron withdrawing nitro group is preferred over the electron releasing methoxy group. This phenomenon can also be seen when examining the *meta*-substituted compounds **3** (*o*-bromo), **4** (*o*-trifluoromethyl) and **11** (3,5-dinitro) with good β-haematin inhibition activities of 24.5 ± 4.2 μM, 28.1 ± 1.7 μM and 13.3 ± 0.5 μM noting that all substituents are electron withdrawing. Compound **11** with the two electron withdrawing nitro groups showed the best β-haematin inhibition activity of the series. Nitro groups are more strongly electron withdrawing than halogens. The conclusions that can be drawn from this series are: electron withdrawing groups (EWG) result in increased β-haematin inhibition activity over electron releasing groups (ERG) and the more electron withdrawing the group, the better the activity. It can also be noted that the order for increasing β-haematin inhibition activity for various substituent positions on the ring system are as follows: *para* < *ortho* < *meta.*

The type of ring system was also shown to influence the β-haematin inhibition activity of these compounds. To analyse this, a comparison between compounds **2** (phenyl) and **16** (pyridyl) was made. The difference between these two compounds can be seen immediately with compound **2** showing complete inactivity and compound **16** showing a moderate β-haematin inhibition activity of 35.6 ± 1.7 μM. This trend is also seen when comparing compounds **4** (*o*-trifluoromethylphenyl) and **9** (*o*-trifluoromethylpyridyl) with respective β-haematin activities of 28.1 ± 1.7 μM and 16.6 ± 0.3 μM. The modification of a 6-membered ring system to a 5-membered ring system showed moderate to poor activity when comparing compounds **17** (thiophene), **18** (furan) and **19** (pyrrole) with β-haematin inhibition activities of 42.4 ± 2.1 μM, 84.1 ± 2.4 μM and 46.6 ± 1.7 μM respectively to compound **16** with an electron poor pyridyl ring with a β-haematin inhibition activity of 35.6 ± 1.7 μM. However, these modifications are all superior to the electron rich phenyl ring system with compound **2** being inactive (> 100 μM). Thus the order for increasing β-haematin inhibition activity for various ring systems is as follows: phenyl < furan < pyrrole < thiophene < pyridyl. All these findings are summarised in Scheme 4.3.



Scheme 4.3: The SAR trends found for the benzimidazole series for β-haematin inhibition with good/ preferred groups in green, moderate groups in blue and poor groups in red.

4.3.3 Molecular Docking

To gain more insight into the observed SARs, molecular docking was attempted. Molecular docking is the process whereby a molecule is positioned in an active site in such a way as to try to achieve maximum binding affinity and molecular interactions between the two sites.[158–161] This is usually used to determine ligand-protein interactions but can in principle also be used for interactions with a crystal surface.[158] This model can assist in identifying the key interactions between the ligand and active site for biological activity as well as assist in verifying a proposed mechanism of action for certain drug-like molecules.[158]

Even though one of the first suggestions of molecular docking came from Crick, it was only in the mid-1980s that the field of molecular docking started to develop.[159] The first widely used docking programme was simply called DOCK and was developed by Kuntz *et al*.[159] However, even with the development of docking programmes, the process of docking is not simple. Ideally, one has to consider several enthalpic and entropic factors which play key roles in binding, the mobility of the active site and ligands and interactions with other molecules in solution such as water, although approximations often have to be made.[161] There are three key considerations for any docking study: the representation of the system of interest, the conformational searching algorithm and the scoring function.[159] Three families of scoring

functions exist namely: force-field based, knowledge based and empirical. Similarly, three conformational algorithms exist namely: systematic methods (incremental construction and conformational search), simulation methods (energy minimisation and molecular dynamics) and random methods (Monte Carlo, Tabu search and Genetic Algorithms).[138,160] This makes each docking programme unique and thus different docking programmes will be more efficient for different systems. Solving a docking problem for a certain system requires the use of a fast and efficient searching algorithm coupled with a fast enough scoring function that can easily discriminate between native or non-native conformations of the ligand and binding site and is unique to the system in question.[159]

4.3.3.1 Materials Studio

Of all the docking programmes available including AutoDock, FlexX, FRED, Glide, GOLD and DOCK, the Materials Studio package [162] is the only one specifically designed to be able to compute adsorption calculations with different surfaces. The package is able to handle crystal surfaces with promising results.[138] This makes it ideal for docking studies using a series of small molecules adsorbed onto a β-haematin crystal structure. The Materials Studio package contains many functions but for this work only the Forcite geometry optimisation, adsorption locator with simulated annealing (Forcite anneal) functions and energy calculations were necessary.

A Forcite geometry optimisation is a single point energy calculation that determines the minimum energy of the molecule in question. The adsorption locator uses an input substrate (in this case β-haematin crystal structure) and an input adsorbent (benzimidazole). The simulation uses the Monte Carlo method to determine the optimum orientation of the absorbent on the substrate.[163] The calculation also uses simulated annealing to find a minimum energy structure. These two techniques are simultaneously used through multiple cycles to allow for more accurate results.[163] Annealing is a technique where the molecule in question is heated to high temperatures and subsequently cooled over cycles to find the optimum configuration.[163] As with all computational systems there are limitations using this methodology for docking. The most noteworthy of these limitations is the exclusion of solvent during calculations (calculations are carried out in a vacuum slab). This can be problematic in terms of direct comparison to working under physiological conditions, as water can interact with both the molecule and β-haematin crystal surface changing the outcome of

intermolecular interactions. This is however unavoidable as including water in these simulations would increase the computational time to the extent of making the simulation impractical. This is exacerbated by having to investigate different crystal faces for each compound. The other limitation to this methodology is that the entropy of the system is not considered as only the energy of the β-haematin crystal and compound are investigated. In order to investigate this entropy change would require more sophisticated techniques adding to time and resource costs.

4.4.3 Docking Results and Discussion

Confirmation of the benzimidazole compounds as β-haematin inhibitors led to the investigation of the binding configurations and energies on the fastest and second fastest growing β-haematin crystal faces,[84] (001) and (011), using molecular docking techniques. Materials Studio was used for all docking computations which allowed the use of a β-haematin crystal surface based on the reported crystal structure[84] and modified CVFF force-field[84] produced by the de Villiers group (unpublished work).

Compounds **2**-**19** were drawn in Chemdraw Ultra 12.0.2 and imported into Materials Studio where the geometry of the compounds were optimised using a Forcite geometry optimisation calculation. These optimised compounds were adsorbed onto both of the pre-built β-haematin crystal surfaces using the adsorption locator. The lowest energy adsorption results for each compound and crystal face were subjected to a Forcite annealing calculation to obtain the lowest plausible energy configuration of the system. An energy calculation was performed on this β-haematin crystal and compound complex. This was used to obtain the adsorption energy for each compound on each crystal face using equation 4.1:

$$E_{ads}\ (kcal.mol^{-1}) = E_{final} - E_{initial}$$

$$= E\ (crystal + inhibitor) - E\ (crystal) - E\ (inhibitor) \qquad \textbf{Equation 4.1}$$

The adsorption results are shown in Table 4.3 alongside the experimental β-haematin inhibition results discussed earlier.

Table 4.3: The adsorption energies obtained from the molecular docking on both β-haematin crystal faces of compounds **2**-**19.**

Compound	β-haematin inhibition IC_{50} (μM)	Adsorption energy (kcal/mol) (001)	Adsorption energy (kcal/mol) (011)
2	Inactive	-59.80	-57.23
3	24.5 ± 4.2	-61.35	-62.20
4	28.1 ± 1.7	-68.73	-69.95
5	32.0 ± 1.4	-60.73	-61.67
6	Inactive	-61.24	-60.26
7	Inactive	-68.21	-64.62
8	38.9 ± 1.5	-60.98	-60.42
9	16.6 ± 0.3	-64.60	-68.39
10	28.9 ± 1.6	-71.45	-64.34
11	13.3 ± 0.5	-84.11	-74.30
12	16.8 ± 0.4	-66.40	-66.41
13	38.2 ± 1.4	-64.62	-62.35
14	25.1 ± 1.1	-63.79	-62.62
15	Inactive	-66.83	-64.58
16	35.6 ± 1.7	-60.59	-58.62
17	42.4 ± 2.1	-60.84	-58.48
18	84.1 ± 2.4	-61.71	-58.08
19	46.6 ± 1.7	-61.33	-57.49

Analysis of this data lead to a simple correlation of the β-haematin inhibition IC_{50} values with the adsorption energy on the (011) β-haematin crystal face, shown in Figure 4.3. This suggests that the (011) β-haematin crystal face was the most important crystal face for the

benzimidazole compounds to adsorb onto in order to inhibit β-haematin formation and therefore this was the topic for further analysis. In order to justify this correlation, an in-depth analysis of the intermolecular interactions between this series of compounds and β-haematin was performed.

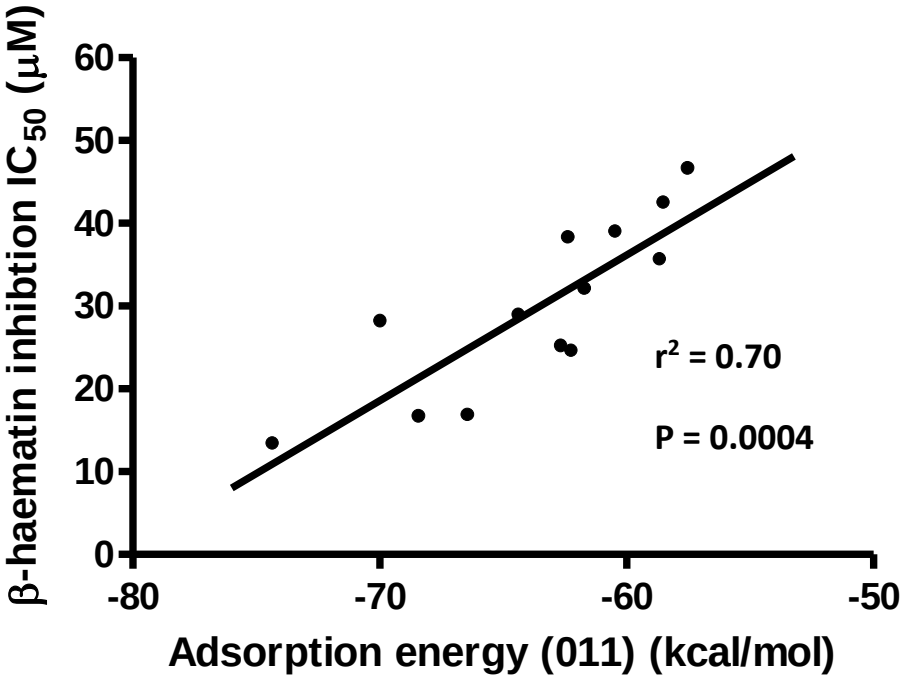
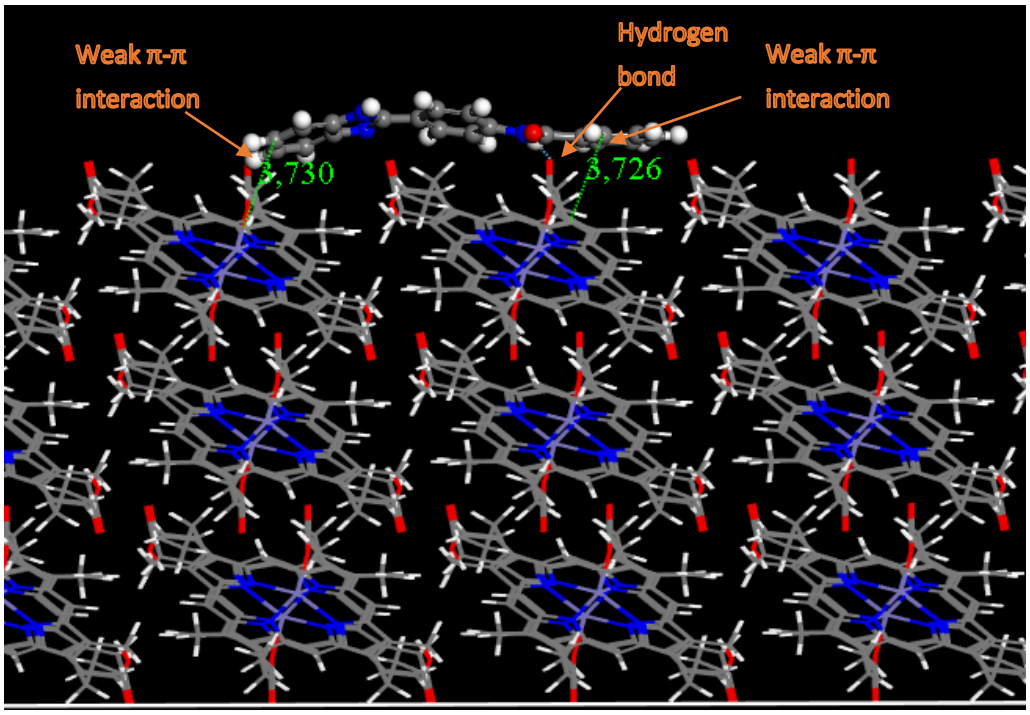


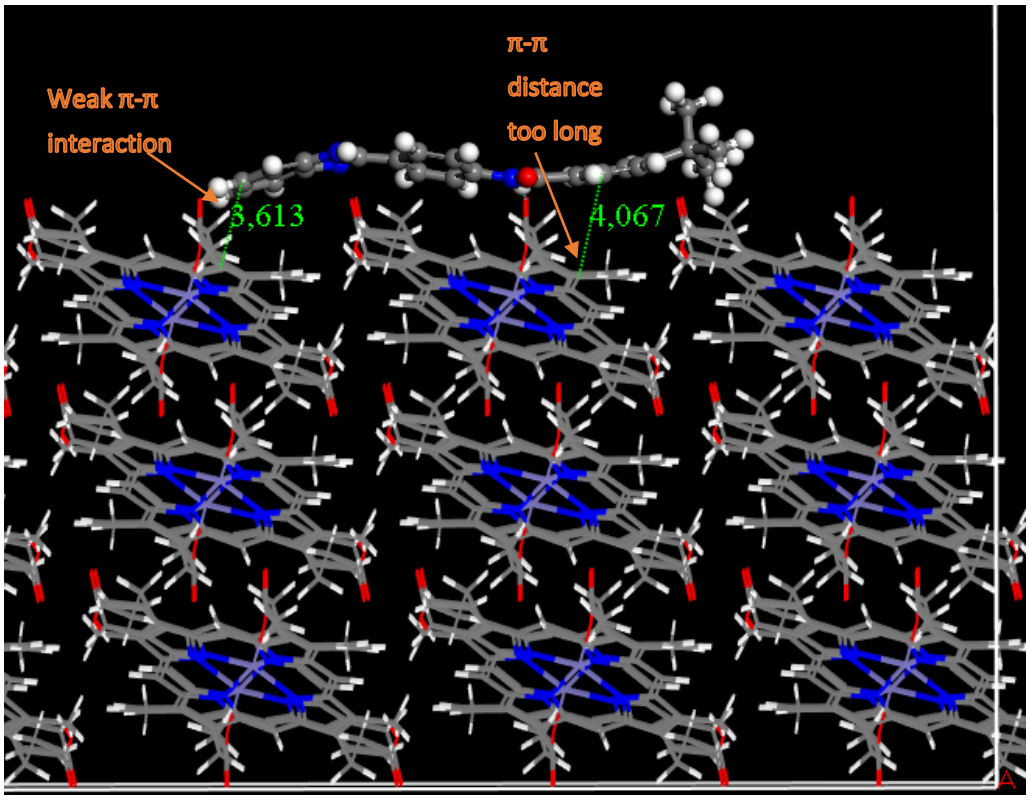
Figure 4.3: Experimental β-haematin IC_{50} versus the adsorption energy for the (011) β-haematin crystal face.

Using intermolecular interaction distances for hydrogen bonding, which were taken to be optimally between 2.36 and 3.69 Å,[164] and π-π interactions distances which were optimally taken to be between 3.3 and 3.8 Å, [165] molecular docking results could be rationalised.

Examining the predicted structures of the experimental β-haematin non-inhibitors **2, 6, 7** and **15** it was seen that these compounds showed π-π interactions as well as hydrogen bonding with the β-haematin crystal surface (Figure 4.4). Compound **2** showed two π-π interactions both with distances of 3.7 Å. Compared to literature standards this is near the upper limit and thus termed weak. Compound **2** also contained a short hydrogen bond between the amide hydrogen and a propionate group of the (011) β-haematin crystal face with a bond distance of 2.68 Å. Compound **6** contained one π-π intermolecular interaction and a short hydrogen bond with respective distances of 3.6 Å and 2.20 Å. The other π-π distance was found to be

much greater than the upper limit from literature at 4.1 Å. Compound **7** was found to be similar to compound **6** with one π-π interaction with a distance of 3.6 Å and a short hydrogen bond with a bond distance of 2.37 Å. Compound **15** showed two very weak π-π interactions with distances of 3.8 Å which is the upper limit according to literature. This compound also showed a short hydrogen bond with a bond distance of 2.12 Å.





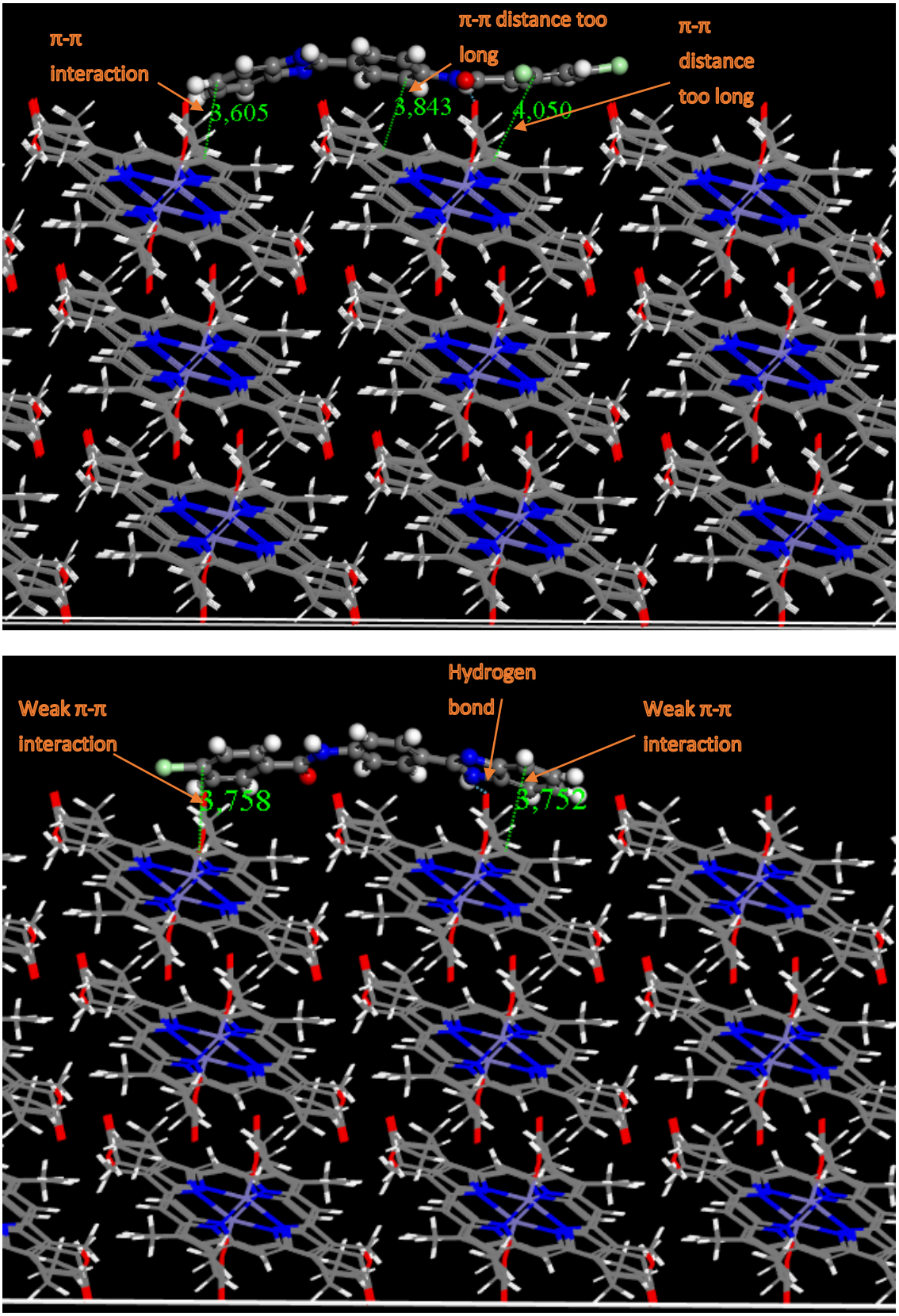


Figure 4.4: Docked structures of compounds **2, 6, 7** and **15** (top to bottom) on the (011) crystal face of the β-haematin crystal with π-π interactions shown in green and hydrogen bonding shown in light blue.

A similar behaviour compared to the inactive compounds was found with compound **18** (Figure 4.5)**.** This compound showed only one π-π interaction and one short hydrogen bond with respective bond distances of 3.6 Å and 2.41 Å. This is not surprising as the β-haematin inhibition IC_{50} of this compound was found to be 84.1 ± 2.4 µM which is close to the cut-off value of 100 µM.

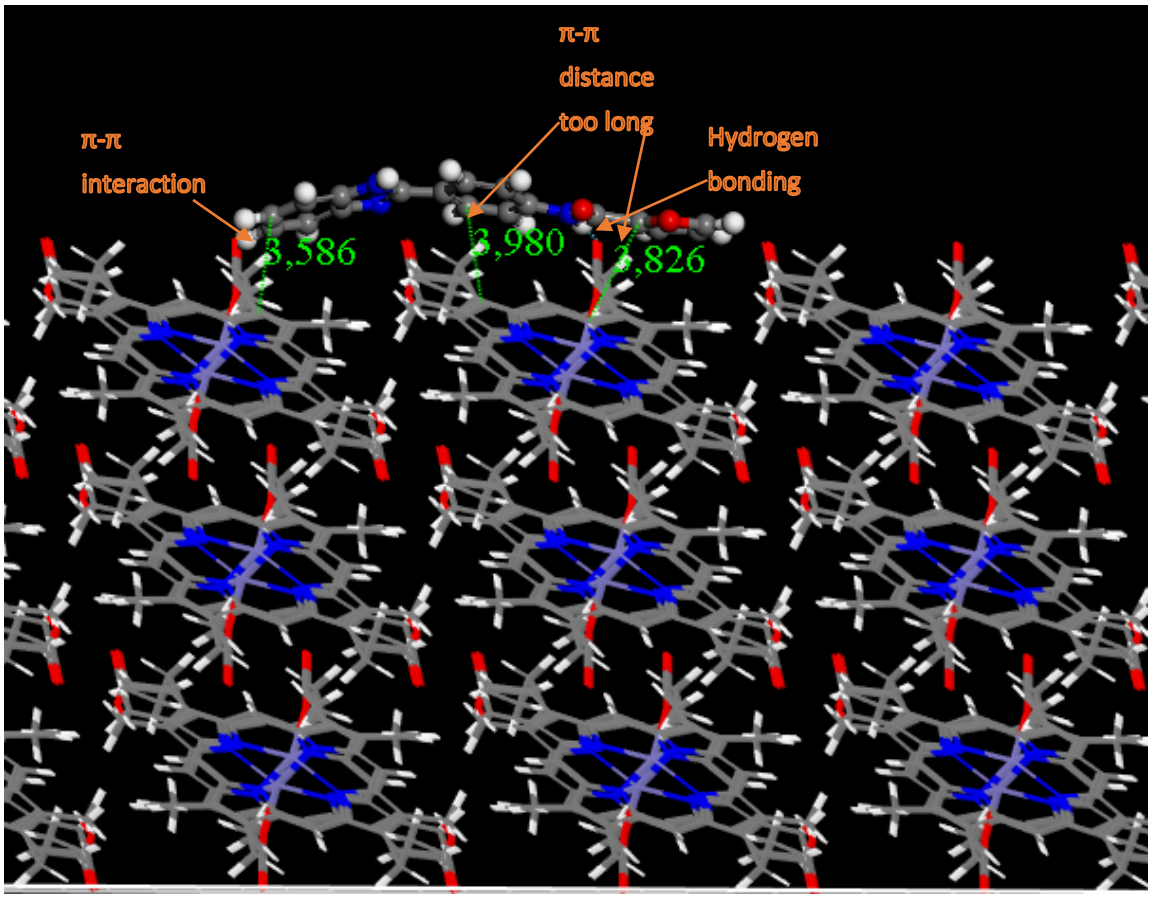


Figure 4.5: Docked structure of compounds **18** on the (011) crystal face of the β-haematin crystal with π-π interactions shown in green and hydrogen bonding shown in light blue.

All other active compounds in this series showed similar behaviour with one to two π-π interactions and a hydrogen bond. Two examples, compounds **3** and **5,** are shown in Figure 4.6 to illustrate this observation. Compound **3** showed two π-π interactions with bond distances of 3.8 Å and 3.7 Å with a moderate hydrogen bond with bond distance of 2.98 Å. Similarly, two π-π interactions are shown with compound **5** with both bond distances of 3.7 Å and one moderate hydrogen bond with bond distance of 2.94 Å.

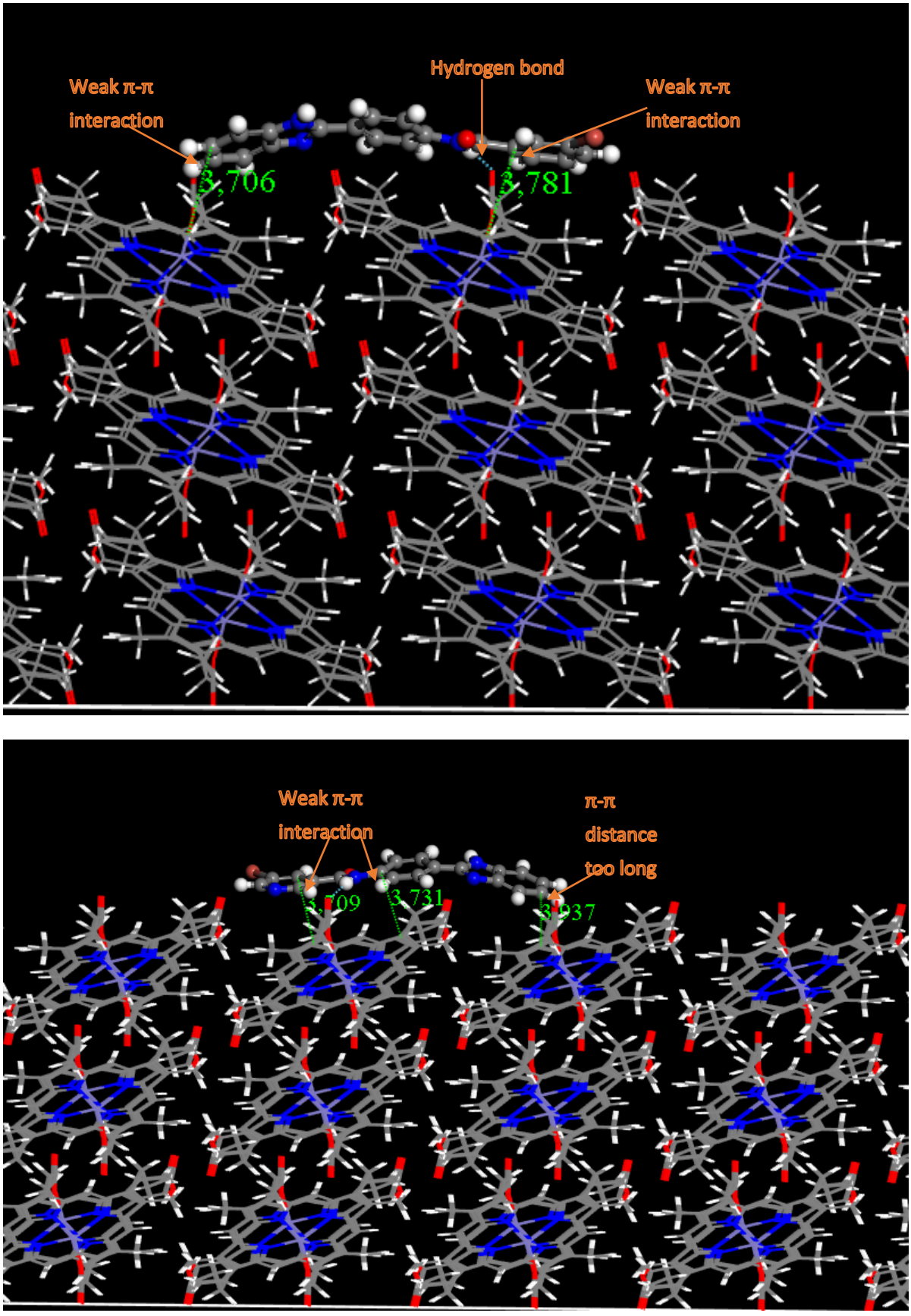
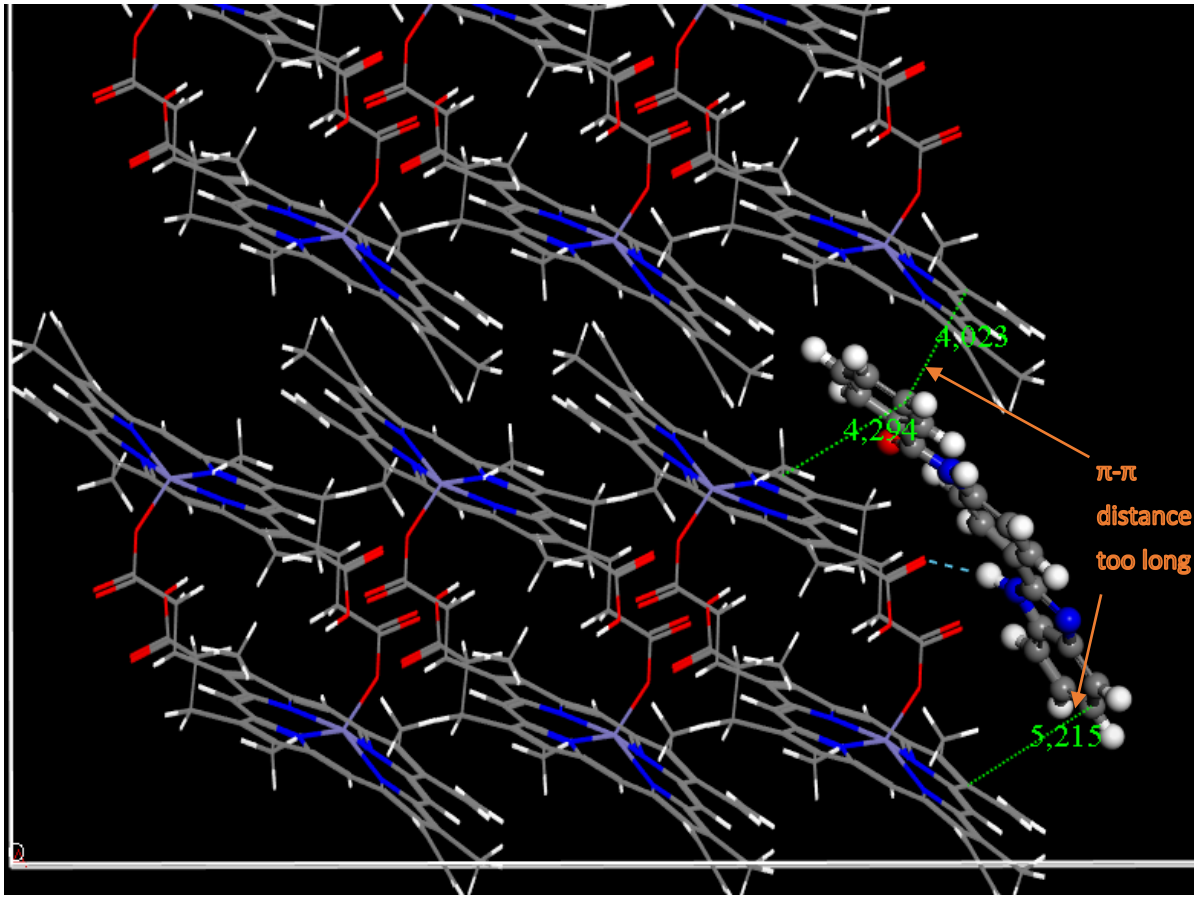
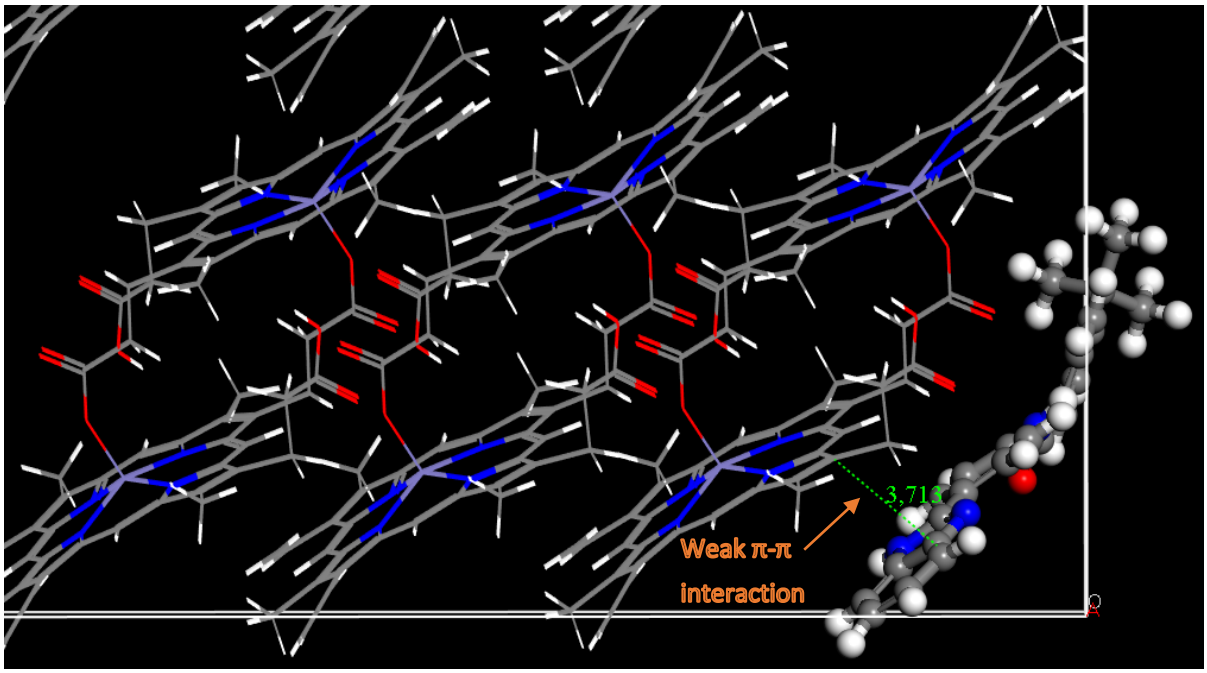


Figure 4.6: Docked structures of compounds **3** and **5** (top to bottom) on the (011) crystal face of the β-haematin crystal.

A clear trend could not be seen between the β-haematin inhibition inactive compounds and β-haematin inhibitors in the series. This led to consideration of the same compounds adsorbed on the (001) β-haematin crystal face even though it was known from the previous correlation that the (011) β-haematin crystal face was the most important in terms of these adsorption energies.

Examining the predicted structures of the experimental β-haematin inhibition inactive compounds **2, 6, 7** and **15**, it was seen that these compounds showed little to no intermolecular interactions with the (001) β-haematin crystal surface (Figure 4.7). Compound **2** showed no π-π interactions as the distances of 5.2, 4.3 and 4.0 Å were much greater than the maximum π-π interaction distances of 3.8 Å. This compound did however contain an anomalously close hydrogen bonding contact with a bond distance of 2.0 Å which is lower than the lower limit for hydrogen bonding of 2.36 Å. In contrast, compound **6** had no hydrogen bonds but did contain one weak π-π interaction between the benzimidazole ring and the porphyrin ring system with a distance of 3.7 Å compared to the maximum distance of 3.8 Å according to literature. Compound **7** contained one intermolecular interaction with the β-haematin surface in the form of a short hydrogen bond with a value of 2.49 Å. This interaction was near the lower limit of literature with a value of 2.36 Å. The π-π interaction indicated in Figure 4.6 of compound **7** could be considered very weak as it falls on the upper limit of 3.8 Å. Compound **15** contains one weak π-π intermolecular interaction between the *p*-chloro substituted ring system and the porphyrin ring system with a distance of 3.5 Å. This compound also contains a short hydrogen bond with the bond distance of 2.8 Å.


4,023
4,294
π-π
distance
too long
5,215


3,713
Weak π-π
interaction

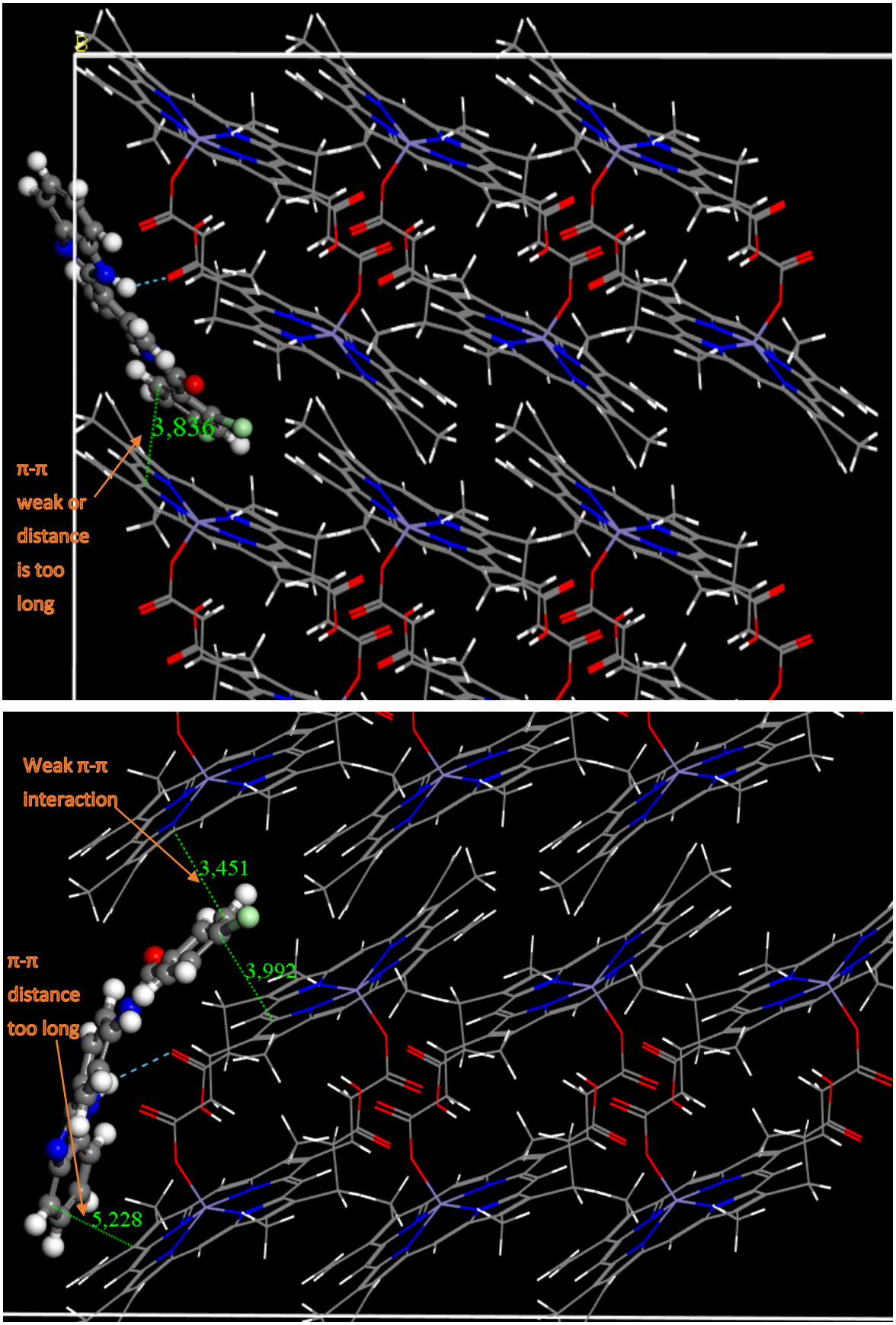


Figure 4.7: Docked structures of compounds **2, 6, 7** and **15** (top to bottom) on the (001) crystal face of the β-haematin crystal showing few to no interactions with the β-haematin surface.

Compound **18** (Figure 4.8) showed only one π-π interaction with a interaction distance of 3.7 Å and no hydrogen bonding. This supported the assumption that **18** was too weak a β-haematin inhibitor to be included in the correlation study as it contained the same number of π-π interactions as the inactive compounds which were not included in the previous correlation.

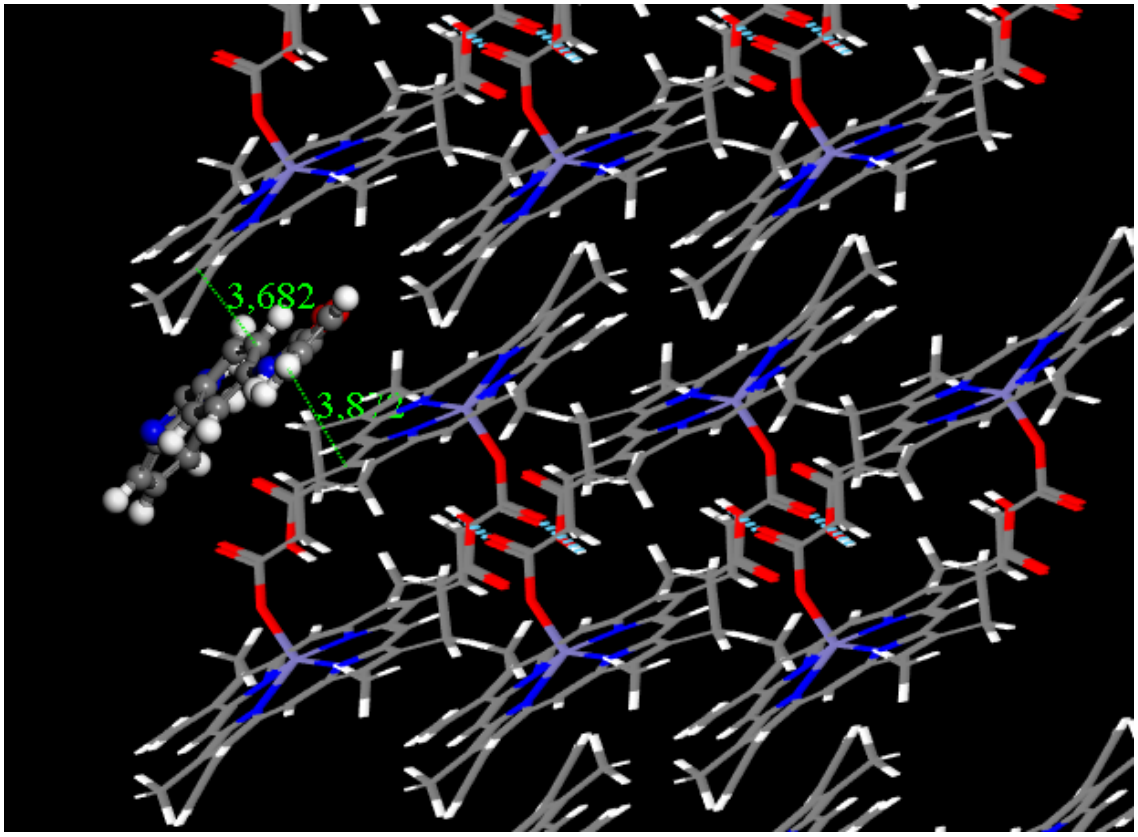


Figure 4.8: Docked structures of compounds **18** on the (001) crystal face of the β-haematin crystal.

By contrast the experimentally active β-haematin inhibitors contained between two and three moderate to strong intermolecular interactions with the β-haematin crystal surface (Figure 4.9). Two π-π interactions were visualised between the *m*-bromo-substituted ring system and the porphyrin ring system as well as the benzimidazole ring system and the porphyrin ring system with respective bond distances of 3.8 and 3.5 Å for compound **3.** Compound **3** also contained a short hydrogen bond with a bond distance of 2.19 Å. Similarly, compound **5** contained two π-π interactions with distances of 3.5 and 3.7 Å and a hydrogen bond with bond distance of 2.17 Å.

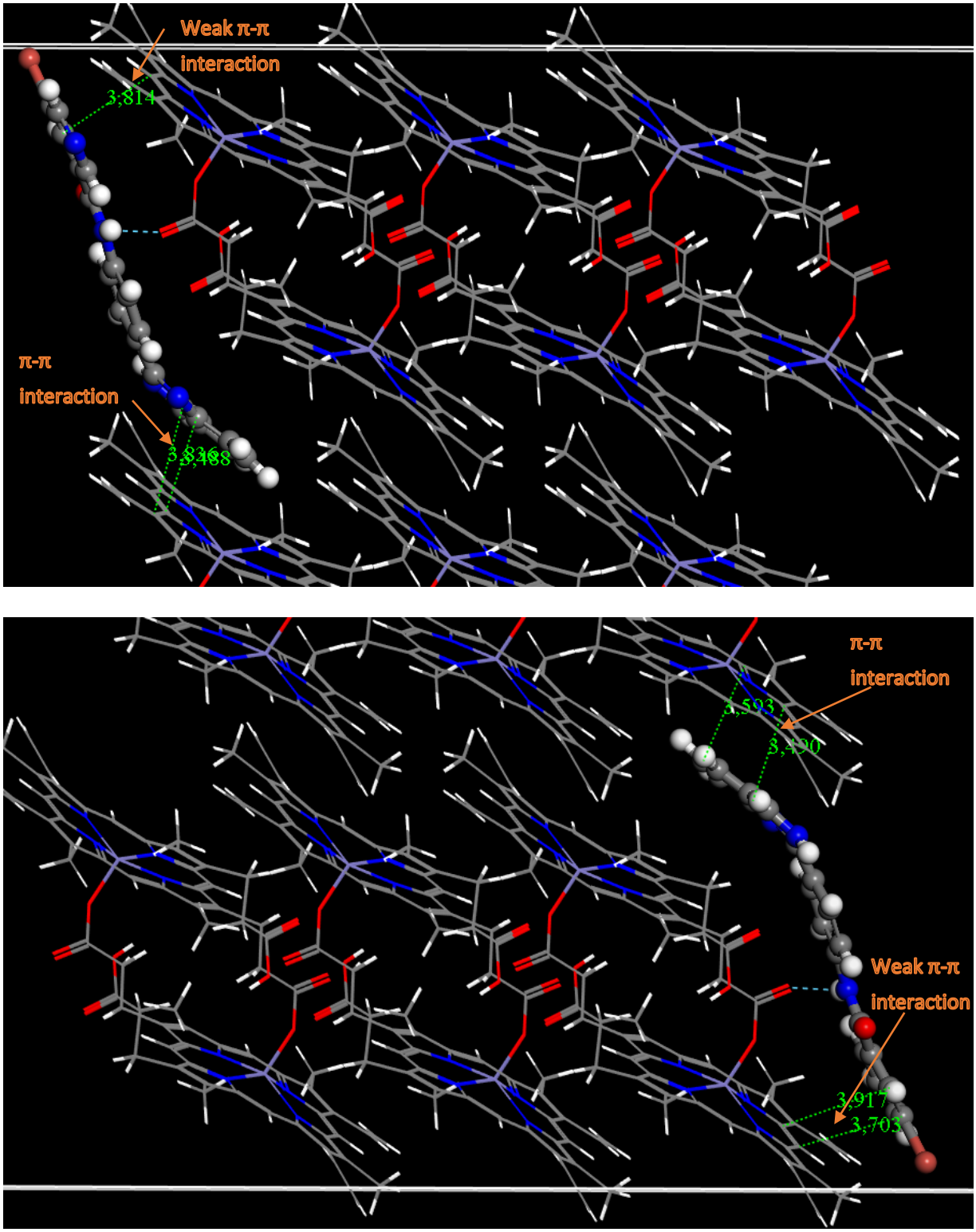


Figure 4.9: Docked structures of compounds **3** and **5** (top to bottom) on the (001) crystal face of β-haematin showing more favoured intermolecular interactions of at least two π-π interactions and a hydrogen bond per molecule.

A summary of the π-π interactions are shown in Table 4.4. This indicated that any compound with two or less π-π interactions on a combination of both β-haematin crystal faces showed inactivity. The only outlier to this is compound **15** which showed three π-π interactions, however those interactions are extremely weak as both interactions on the (011) crystal face were found to have values of 3.8 Å which is the upper limit according to literature. These observations justified the exclusion of compounds **2**, **6**, **7**, **15** and **18** from the correlation of adsorption energies of the (011) β-haematin crystal face and the β-haematin inhibition IC_{50} values on the basis that a minimum number of specific interactions are also required.

Table 4.4: The number of π-π interactions for each of the β-haematin crystal faces of compounds **2-19** with inactive compounds and outliers shown in red.

Compound	No. of π-π interactions on (011) β-haematin crystal face	No. of π-π interactions on (001) β-haematin crystal face	Total No. of π-π interactions
2	2	0	2
3	2	2	4
4	1	2	3
5	2	2	4
6	1	1	2
7	1	0	1
8	1	2	3
9	2	2	4
10	2	1	3
11	2	1	3
12	2	1	3
13	2	1	3
14	2	2	4
15	2	1	3
16	2	1	3
17	2	1	3
18	1	1	2
19	2	1	3

4.4.4 Malaria Parasite Growth Inhibition Studies

Following the investigations into the β-haematin inhibition of the benzimidazole compounds using the various tools described above, compounds **2**-**19** were tested for parasite activity *in vitro*. The malaria parasite growth inhibition activity assay was carried out by collaborators using the lactate dehydrogenase assay instead of the SyBr Green assay as previously discussed, which used cultured NF54 chloroquine sensitive malaria parasites and were tested in triplicate.[166,167] The results are shown in Table 4.5 alongside the predicted Bayesian score.

Of the eighteen predicted inhibitors of malaria parasite growth, only two were found to be experimentally inactive using a cut-off of 20 µM. This meant that 89% of the compounds were predicted correctly by the Bayesian model corresponding to a 45-fold enrichment compared to random screening using the malaria parasite growth inhibition Bayesian model containing 817 active compounds and 41 729 inactive compounds.[127] If, however, a cut-off of 2 µM (used because the compounds were predicted using a pre-built Bayesian model) was used, as was the case with the Bayesian predictions, nine out of the eighteen compounds would be termed active and with a hit rate of 50%. When compared to random screening using the same Bayesian model a 26-fold enrichment was achieved. This false positive data was generated because the Bayesian model recognises any compound with a benzimidazole fingerprint as active, so with the new input of this data the Bayesian model would be improved. Thus, not only was the Bayesian screening highly successful using the β-haematin inhibition model, but also for the malaria parasite growth inhibition model.

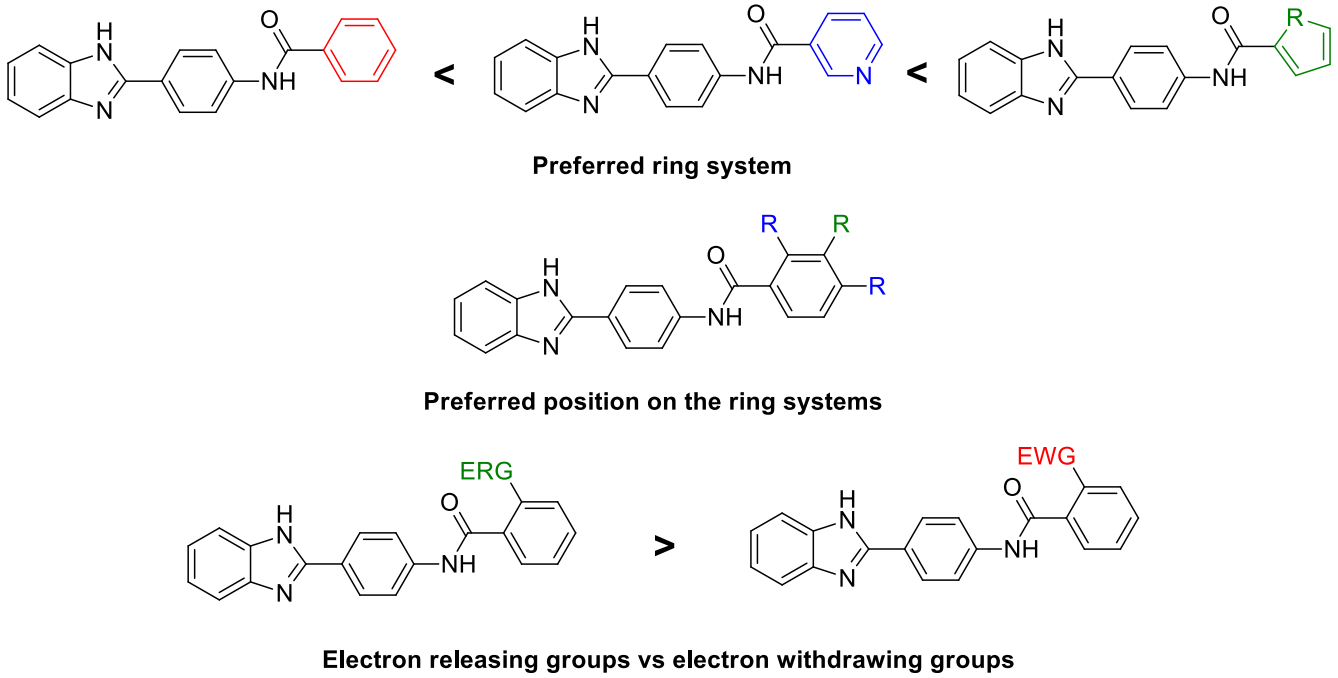
Table 4.5: The malaria parasite growth inhibition IC_{50} results of compounds **2**-**19** using the NF54 strain of *P. falciparum* and the accompanying Bayesian prediction from Chapter 2.

Compound	NF54 malaria parasite growth inhibition IC_{50} (μM)	Malaria parasite growth inhibition score
2	Inactive	61.3259
3	1.3 ± 0.2	57.2935
4	1.34 ± 0.07	75.2685
5	0.41 ± 0.03	72.6119
6	2.0 ± 0.1	63.8076
7	7.2 ± 0.2	49.3979
8	5.50 ± 0.6	43.3265
9	0.8 ± 0.5	81.8491
10	12.3 ± 2.7	30.0400
11	0.67 ± 0.07	47.2409
12	0.53 ± 0.13	65.4908
13	15.0 ± 0.3	45.2802
14	1.21 ± 0.06	58.6210
15	Inactive	56.5011
16	11.6 ± 2.6	78.3569
17	7.2 ± 0.7	55.1323
18	5.9 ± 0.8	48.2215
19	1.2 ± 0.4	67.9852

4.4.5 Benzimidazole Malaria Parasite Growth Inhibition SAR

As with the β-haematin inhibition SAR, a comparison of **13** (*o*-chloro), **14** (*m*-chloro) and **15** (*p*-chloro) showed a trend for position of substitution. Compound **15** again showed poor activity with a parasite growth inhibition activity of greater than 20 µM, while compound **13** showed moderate activity with an IC_{50} of 15.0 ± 0.3 µM and compound **14** showed good activity with an IC_{50} value of 1.21 ± 0.06 µM. This followed the same trend as the β-haematin inhibition activity of *para* < *ortho* < *meta*. However, upon examination of compound **6** (*p*-*tert*-butyl), an IC_{50} of 2.0 ± 0.1 µM was obtained. This indicated that not all *para*-substituted ring systems showed poor activity. In fact, this biological activity was superior to all *ortho*-substituted compounds in the series. When a comparison between compound **8** (*o*-methoxy) and **10** (*o*-nitro) was made, a conclusion was drawn that the electron releasing groups actually showed an increased activity compared to electron withdrawing groups. Therefore, the final conclusion that was drawn for the position of substitution trend for malaria parasite growth inhibition was as follows: *para* ≈ *ortho* < *meta*.

Upon investigating the optimal type of ring system, a comparison between **2** (phenyl) and **16** (pyridyl) was informative. Compound **2** had an IC_{50} value greater than the 20 µM cut-off while compound **16** had an IC_{50} value of 11.6 ± 2.6 µM. This trend also holds when comparing the substituted compounds **3** (*m*-bromophenyl) and **5** (*m*-bromopyridyl) with respective IC_{50} values of 1.3 ± 0.2 µM and 0.41 ± 0.03 µM. Similarly, this was seen with compounds **4** (*m*-trifluoromethylphenyl) and **9** (*m*- trifluoromethylpyridyl) with respective IC_{50} values of 1.34 ± 0.07 µM and 0.8 ± 0.5 µM respectively. Both these sets of compounds also showed a promising activity shift to the nM range, although their activities are not close to that of the chloroquine standard used (12.51 ± 0.63 nM). When the two six-membered ring system values were compared to the five-membered ring series of compounds **17** (thiophene), **18** (furan) and **19** (pyrrole) with respective IC_{50} values of 7.2 ± 0.7 µM, 5.9 ± 0.8 µM and 1.2 ± 0.4 µM, the most active compounds in this comparison were definitely the 5-membered ring systems. Thus, the activity trend for the differing ring systems when comparing malaria parasite growth inhibition IC_{50} values was as follows: phenyl < pyridyl < thiophene < furan < pyrrole. All these findings are summarised in Scheme 4.4.



Scheme 4.4: The SAR trends found for the malaria parasite growth inhibition IC_{50} of the benzimidazole series with preferred groups in green, moderate groups in blue and poor groups in red.

4.4.6 Parasite Growth Inhibition Studies in the Dd2 Strain and Cytotoxicity

A selection of the most active compounds from the NF54 testing were further tested on the chloroquine resistant strain of malaria parasite, Dd2, using the same assay conditions to determine if cross-resistance was evident for this series of compounds. The results are shown in Table 4.6 alongside the NF54 results for comparison.

The Dd2 results showed a slight decrease in activity of compounds **5** (*m*-bromo) and **9** (*m*-trifluoromethyl) with IC_{50} values of 0.96 ± 0.04 µM and 1.10 ± 0.37 µM respectively compared to the NF54 results of 0.41 ± 0.03 µM and 0.8 ± 0.5 µM. The opposite was true for compounds **11** (3,5-dinitro) and **12** (2,6-dichloro) which showed small increases in activity going from IC_{50} values of 0.67 ± 0.07 µM and 0.53 ± 0.13 µM respectively in the NF54 strain to values of 0.31 ± 0.02 µM and 0.34 ± 0.03 µM respectively in the Dd2 strain. This was clearly seen with the resistance indexes (RIs) of compounds **5** and **9** greater than one and while compounds **11** and **12** showing values of less than one. Using this data it can be seen that for inhibition of the Dd2 strain, a disubstituted compound is preferred to a monosubstituted compound. None of the compounds exhibit cross-resistance, since the RI values are all small.

Table 4.6: The malaria parasite growth inhibition IC_{50} results of compounds **5, 9, 11** and **12** using the Dd2 strain of *P. falciparum* and the accompanying NF54 IC_{50} results.

Compound	Dd2 malaria parasite growth inhibition IC_{50} (μM)	NF54 malaria parasite growth inhibition IC_{50} (μM)	Resistance Index (RI)
5	0.96 ± 0.04	0.41 ± 0.03	2.34
9	1.10 ± 0.37	0.80 ± 0.50	1.38
11	0.31 ± 0.02	0.67 ± 0.07	0.45
12	0.34 ± 0.03	0.53 ± 0.13	0.64

The cytotoxicity of compounds **5**, **9**, **11** and **12** was determined against the Chinese Hamster Ovarian (CHO) cell line, with the results shown in Table 4.7. All compounds were shown to be selective against *P. falciparum* with the selectivity index (SI) ranging from 31.82 to 664.77 in the chloroquine resistant strain and 43.75 to 461.24 in the chloroquine sensitive strain.

Table 4.7: The cytotoxicity results from the CHO cell line of compounds **5, 9, 11** and **12** with the selectivity indexes to the Dd2 and NF54 strain of *P. falciparum*.

Compound	CHO cytotoxicity (μM)	Dd2 malaria parasite growth inhibition IC_{50} (μM)	Selectivity Index (SI) Dd2	NF54 malaria parasite growth inhibition IC_{50} (μM)	Selectivity Index (SI) NF54
5	189.11 ± 4.17	0.96 ± 0.04	196.99	0.41 ± 0.03	461.24
9	35.00 ± 4.39	1.10 ± 0.37	31.82	0.80 ± 0.50	43.75
11	206.08	0.31 ± 0.02	664.77	0.67 ± 0.07	307.58
12	183.87 ± 8.67	0.34 ± 0.03	540.79	0.53 ± 0.13	346.92

4.4.7 Malaria Parasite Growth Inhibition Studies Using a Fragment Approach

The fragment compounds were also tested against the NF54 strain of *P. falciparum* despite their lack of β-haematin inhibition activity. The diacylated product (**20**) was also included in

this study. The results are shown in Table 4.8 along with the Bayesian predictions for the compounds.

All fragment compounds were found to be inactive in comparison to their Bayesian predictions of two inactive compounds. Thus, in the case of this subset of compounds, the Bayesian model predicted 71% of the data correctly with a 36-fold enrichment compared to random screening using the malaria parasite growth inhibition Bayesian model. From this data it can be concluded that the whole molecule is required for biological activity as the removal of any group from the molecule leads to biological inactivity in the malaria parasite. Compound **20** also showed a two-fold diminished malaria parasite growth inhibition activity when compared to its parent compound, **3,** indicating that a substituent on the N atom of the benzimidazole ring is not favoured for parasite activity.

Table 4.8: The malaria parasite growth inhibition IC_{50} results of the fragment compounds and **20** using the NF54 strain of *P. falciparum* accompanied by the Bayesian predictions from Chapter 2.

Compound	NF54 malaria parasite growth inhibition IC_{50} (μM)	Malaria parasite growth inhibition score
1	Inactive	56.9520
20	2.6 ± 0.6	-
21	Inactive	-8.57182
21a	Inactive	-10.6894
22	Inactive	-8.25818
22a	Inactive	-12.8326
23	Inactive	60.8754
23a	Inactive	4.55447

4.4.4 Cellular Haem Fractionation Assay

Following the investigation of the initial benzimidazole series using the NP-40 assay, it was known that the vast majority of these compounds were β-haematin inhibitors. However, this

does not necessarily mean that they will inhibit haemozoin formation in the cell. Without additional data there was uncertainty that the compounds were able to enter the food vacuole. Yet another scenario might be that the compounds had a stronger activity against another target and thus killed the cell in that way without affecting haemozoin formation. In order to confirm that these compounds do indeed inhibit haemozoin formation in the cell, four compounds, **3**, **5**, **6** and **13** were tested by colleagues using the cellular haem fractionation assay developed by Combrinck *et al*.[71,168] This allowed for the measurement of free haem and haemozoin levels in the malaria parasite after dosage at chosen multiples of the IC_{50} of each compound. The results for all the compounds showed similar behaviour to compound **3** shown in Figure 4.10.

The measurements showed that with an increase in concentration of compound **3** there was an evident decrease in both the percentage haemozoin and mass of haemozoin in the malaria parasite as well as a large increase of free haem compared to the control. This is consistent with the proposed mechanism of action being haemozoin inhibition, where more toxic free haem is available to kill the malaria parasite due to the compounds inhibiting the growth of haemozoin.

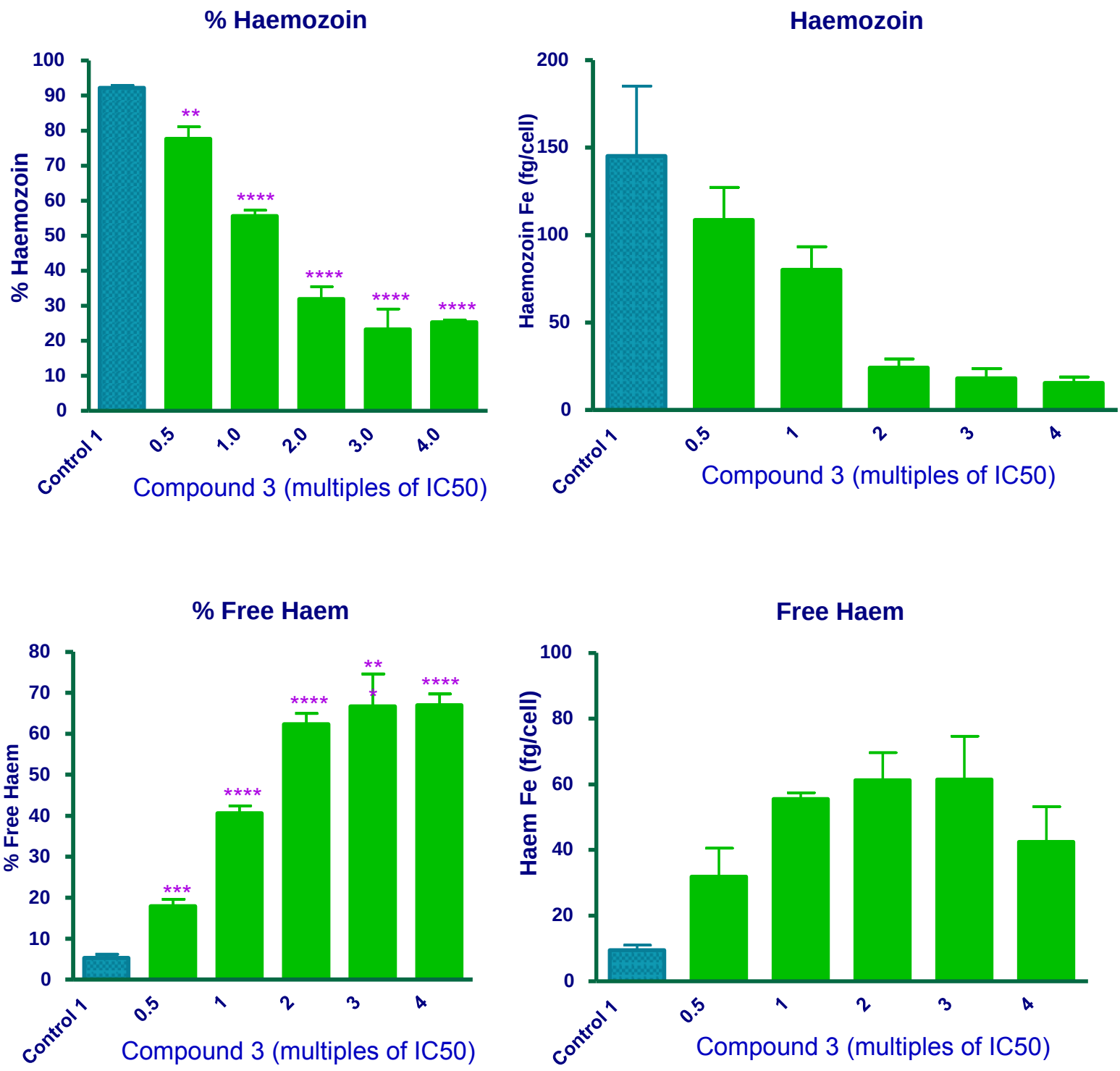


Figure 4.10: The percentage and amount of free haem and haemozoin after incubation with compound **3**.

4.4.5 QSAR Study

A QSAR study on the benzimidazole compounds was conducted using Molecular Modelling Pro Plus (MMPPlus) v 7.0.2.[169] The main input for this study was the experimentally

determined malaria parasite growth and β-haematin inhibition data obtained from the previously discussed studies. Using this and physio-chemical parameters calculated by MMPPlus a feasible multiple correlation was obtained, shown in Figure 4.11.

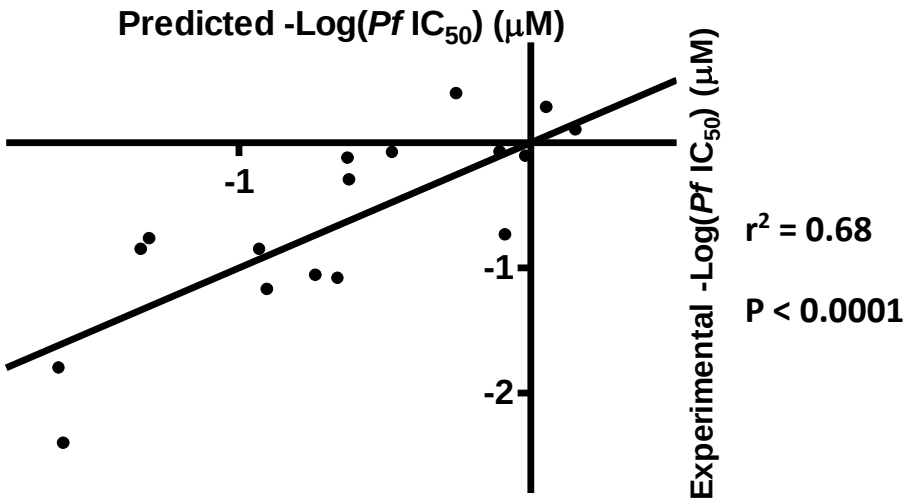


Figure 4.11: Experimental -Log(*Pf* IC_{50}) versus the predicted -Log(*Pf* IC_{50}) obtained from the multiple correlation of various terms according to equation 4.2.

-Log(*Pf* IC_{50}) = (1/β-haematin IC_{50})(28,73) + number of hydrogen bond donors(3,71) + molecular depth(1,43) - 8,90 **Equation 4.2**

This correlation was found to be significant using the F-test at the 99% confidence level (F-stat = 9.84 > F-crit = 5.56) and the two-tailed t-test at 90% confidence level for each parameter (t = 5.12, 1.87, 2.56, 3.46 > t-crit = 1.76). The multiple correlation showed that for an increase in *P. falciparum* activity there was an increase in β-haematin inhibition activity, an increase in the number of hydrogen bond donors and an increase in the molecular depth of the molecule (the amount of space the molecule takes up in terms of height). Looking at the terms individually, in order to obtain a more active compound against malaria parasite growth, the compound was required to have hydrogen bond donors. As shown by the haem fractionation study, the biological activity is linked to their ability to inhibit β-haematin formation. The molecular depth of these compounds remains quite consistent and is the least influential term in the equation. This equation was found to be dominated by the 1/β-haematin IC_{50} which signifies the importance that β-haematin inhibitory activity plays in

this set of compounds and cements the understanding that these compounds act via haemozoin inhibition.

4.5 Summary and Conclusions

In this study all the synthesized benzimidazole compounds were tested using the NP-40 detergent mediated β-haematin inhibition assay. The initial series showed promising results with fourteen out of eighteen compounds termed active. Comparing this to the Bayesian predictions, 78% of the compounds were correctly predicted with a 24-fold enrichment compared to random screening. SAR of the benzimidazole scaffold in terms of β-haematin inhibition showed three trends. A preferred *meta*-substitution pattern was seen and electron withdrawing groups were preferred over electron releasing groups. There was also a preference over which ring system was favoured for β-haematin inhibition activity as follows: phenyl < furan < pyrrole < thiophene < pyridyl. All compounds in the fragment study were found to be inactive in contrast to the Bayesian predictions where only one out of seven predictions was correct, leading to a hit rate of 14% with a 4-fold enrichment compared to random screening.

Docking studies were carried out on all the benzimidazole compounds to investigate the correlation between β-haematin inhibition activity and adsorption energies. A correlation between β-haematin inhibition activity and adsorption energies on the (011) β-haematin crystal face was found. This indicated that the (011) β-haematin crystal face was the most important in terms of adsorption for this series of compounds with respect to inhibition. In order to justify this correlation, the intermolecular interactions of the compounds with the both β-haematin crystal faces was studied. This showed that compounds with less than three π-π interactions were deemed inactive. Compounds **2, 6, 7, 15** and **18** all showed this trend. Hydrogen bonding was also a factor considered in this study but could not explain inactivity as all compounds contained at least one hydrogen bond.

P. falciparum growth inhibition studies were carried out on all compounds using the chloroquine sensitive NF54 strain with selected active compounds also tested on the chloroquine resistant Dd2 strain and for cytotoxicity. Using a cut-off of 20 μM, sixteen out of the eighteen initial benzimidazole compounds were deemed active with 89% (16/18 compounds) of the compounds correctly predicted by the Bayesian statistics leading to a

45-fold enrichment compared to random screening. However, if the cut-off was lowered to 2 µM, as with the initial Bayesian model, only 50% (9/18 compounds) of the compounds were correctly predicted still with a substantial 26-fold enrichment compared to random screening. The results from the NF54 strain showed that *meta*-substituted compounds were favoured and both *para*-substituted and *ortho*-substituted compounds showed lower activity. Electron releasing substituents were favoured in contrast to the finding from the β-haematin inhibition studies. The trend for selection of ring systems was found to be as follows: phenyl < pyridyl < thiophene < furan < pyrrole.

The Dd2 results showed a slight decrease in activity in the monosubstituted compounds while the disubstituted compounds showed a small increase in activity. None showed evidence of cross-resistance with chloroquine. The fragment compounds were all shown to be inactive against the NF54 strain with 71% (5/7 compounds) correctly predicted by Bayesian statistics showing a 36-fold enrichment compared to random screening. Compound **20** saw a 2-fold diminished malaria parasite growth inhibition activity compared to its parent, **3**.

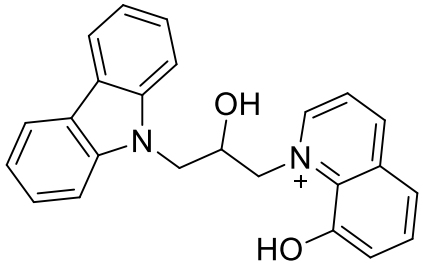
The haem cellular fractionation assay confirmed that the benzimidazole compounds were indeed haemozoin inhibitors and were able to get into the cell and food vacuole with a decrease in haemozoin and an increase in free haem shown. A QSAR study showed that the *P. falciparum* growth inhibition activity values of these benzimidazole compounds can be predicted using the number of hydrogen bond donors, the molecular depth of the molecule and most importantly the β-haematin inhibition activity.

Using the Bayesian model to pre-screen compounds was found to be extremely successful and enabled the quick selection of reliable hit compounds and derivatives thereof. Although the benzimidazole compounds did not show a dramatic improvement in biological activity compared to the hit compound, the findings are none the less important. It is now possible to predict the β-haematin inhibition values by using simulated adsorption energies on the (011) β-haematin crystal face. With these calculated β-haematin values, predictions can be made for parasite activities of new compounds before synthesis. In this way the tedious synthetic steps are mitigated until worthy candidates are discovered using computational techniques.

Chapter 5: Carbazole Indole Molecular Docking and Structure-Activity Relationships for β-haematin and Parasite Growth Inhibition

5.1 Introduction

The initial carbazole scaffold was first discovered in the Vanderbilt high-throughput screen and showed promising biological results (Figure 5.1). However, due to the unfavourable positive charge on the quinoline nitrogen, modifications were made to this scaffold as discussed in Chapter 1. The seven synthesized carbazole indole compounds were investigated for β-haematin inhibition, β-haematin adsorption energies using molecular docking, aqueous solubility and biological activity.



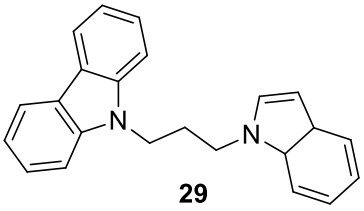
β-Haematin IC_{50} 16.2 µM

Parasite Growth inhibition IC_{50}: D6 200 nM , C235 180 nM

Figure 5.1: The biological results for the hit carbazole compound from the Vanderbilt high-throughput screen.

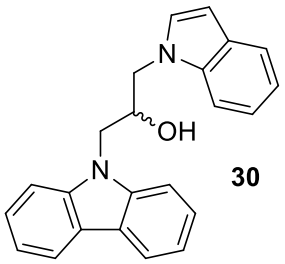
The main focus point for this series of compounds was to investigate how the β-haematin inhibition, biological data and adsorption energies changed by altering the hydrogen bonding capabilities of the molecules through sequential removal of hydroxyl groups from the hit compound analogue (Scheme 5.1). However, as this set of compounds contain a chiral centre all enantiomers needed to be analysed to determine which, if any were active compounds.

Compound with no hydroxyl groups

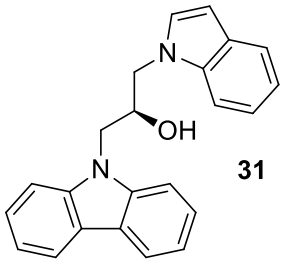


29

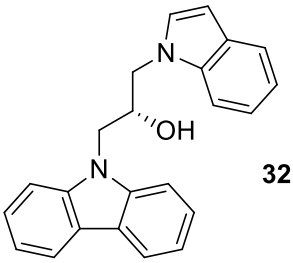
Compounds with one hydroxyl group



30

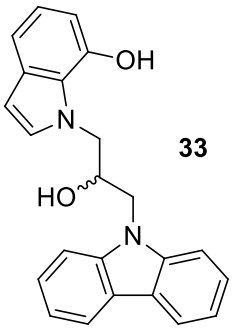


31

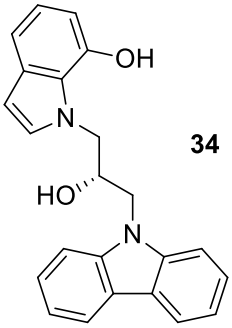


32

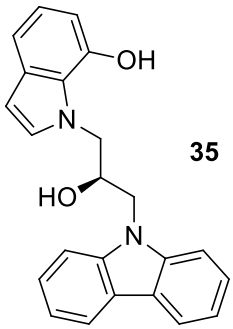
Compounds with two hydroxyl groups



33



34



35

Scheme 5.1: The Synthesized carbazole indole compounds **29-35** with differing number of hydroxyl groups investigated in this chapter.

5.2 Carbazole Indole Molecular Docking and Structure-Activity Relationships Goals

The goals of this chapter were to:

- Perform the NP-40 β-haematin inhibition assay on all seven carbazole indole compounds and compare the results to the Bayesian statistics predictions.
- Subject all seven carbazole indole compounds to parasite growth inhibition tests on the NF54 and the DD2 strains and compare the results to the Bayesian statistics predictions.

- Use Materials Studio to perform docking studies on all of carbazole indole compounds and investigate interactions with the β-haematin crystal surfaces.
- Subject two compounds to the haem fractionation assay.
- Subject two compounds to the turbidimetric solubility assay to test for aqueous solubility.
- Using the data from the above studies, attempt to find a QSAR for compounds **29-35**.

5.3 Results and Discussion

5.3.1 Turbidimetric Solubility Assay

Two of the most important features of compounds in terms of medicinal chemistry are their solubility and pKa values.[170] The measurement of solubility ranges can be performed using a turbidimetric solubility assay. This assay requires a 96 well plate where half the wells incorporate a compound at various concentrations in DMSO as a solvent to act as a baseline reading. Similarly, the other half of the 96 well plate incorporates a compound at various concentrations using a phosphate buffered saline solution as a solvent. In this way the turbidity of the wells can be measured using a UV-visible spectrophotometer at a longer wavelength than usual for organic molecules, 620 nm, to prevent interference from UV active compounds.[171] This is possible as precipitation in the wells can be measured due to the increased light scattering from the precipitated molecules in the well, thus increasing the absorbance value of the well.[172] According to literature, compounds with a solubility of less than 1 µM are conventionally termed highly insoluble, those with solubility values between 1 and 100 µM are termed moderately soluble and compounds with solubility values greater than 100 µM are termed highly soluble.[171]

Thus, the turbidimetric solubility assay was performed on compounds **29**, **32** and **35** in order to determine the aqueous solubility ranges of the compounds. This assay was not performed on the benzimidazole compounds as it was known that all compounds in the benzimidazole series showed solubility ranges of less than 1 µM from other assays performed on those compounds and observations during the synthesis of the benzimidazole compounds. The turbidimetric solubility assay curves are shown in Figures 5.2, 5.3, 5.4 and 5.5 for hydrocortisone and compounds **29**, **32** and **35** respectively**.** The x-axis of the graphs has been modified in order to clearly visualise at what concentration ranges turbidity occurs. Thus 1 on the x-axis represents a concentration of 0.0 µM, 2 a concentration of 5.0 µM, 3 a concentration of 10.0 µM, 4 a concentration of 20.0 µM, 5 a concentration of 40.0 µM, 6 a concentration of 80.0 µM, 7 a concentration of 160.0 µM and 8 a concentration of 200.0 µM.

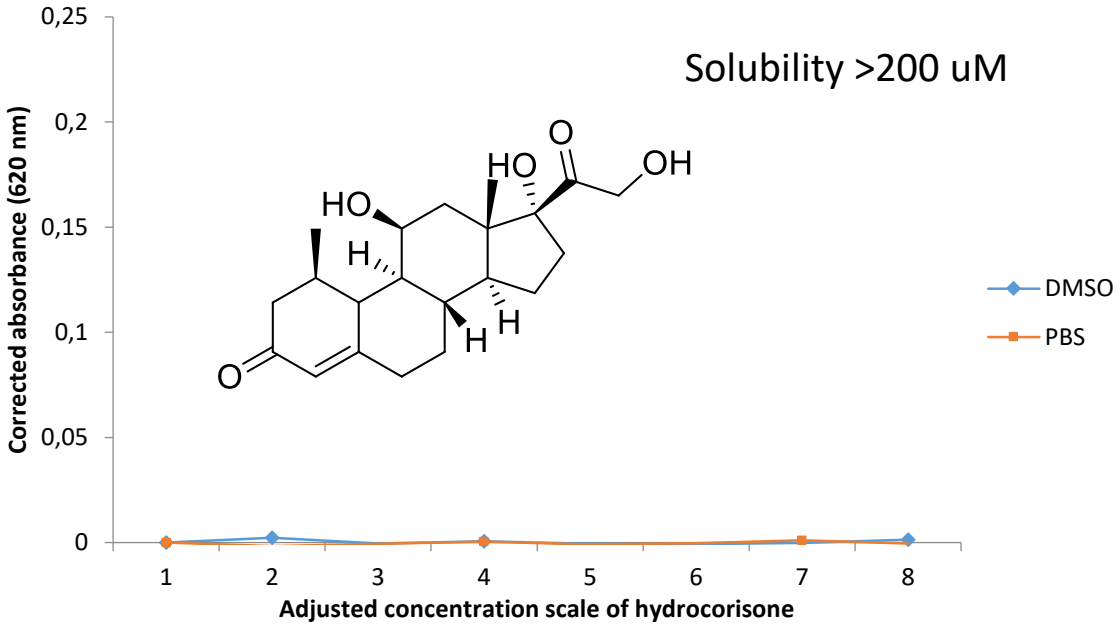


Figure 5.2: Corrected absorbance at 620 nm for hydrocortisone versus the adjusted concentration scale.

Hydrocortisone was used as a positive standard for this assay as it is highly soluble in both DMSO and aqueous medium with a solubility range of greater than 200 µM.

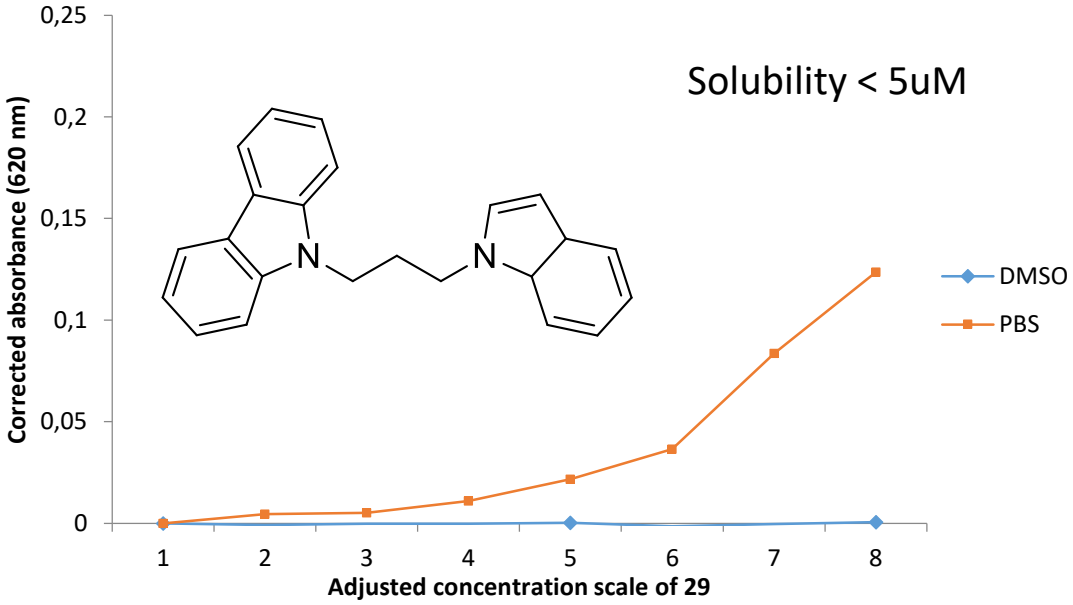


Figure 5.3: Corrected absorbance at 620 nm for **29** versus the adjusted concentration scale.

Compound **29** exhibited a solubility of less than 5 µM as the deviation from the baseline begins between 1 (0.0 µM) and 2 (5.0 µM) on the x-axis. This compound was thus termed

highly insoluble to insoluble as there is no certainty that the concentration is greater than 1 µM.

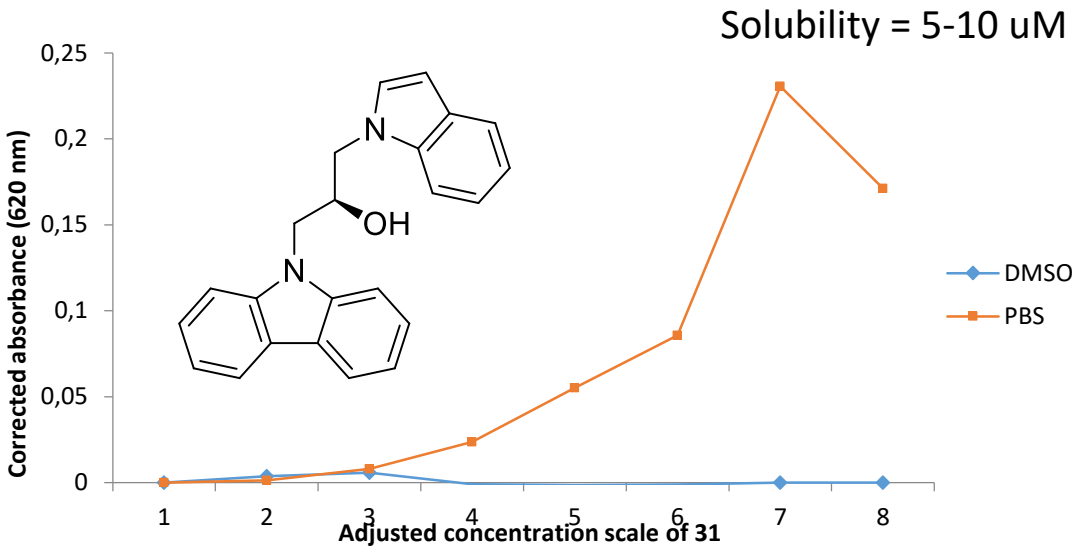


Figure 5.4: Corrected absorbance at 620 nm for **31** versus the adjusted concentration.

Compound **31** exhibited a solubility range of between 5 to 10 µM. This was due to the deviation from the baseline occurring between 2 (5.0 µM) and 3 (10.0 µM) on the x-axis. Therefore, this compound was termed moderately aqueous soluble. This increased solubility was due to the addition of a hydroxyl group on the alkyl chain of the molecule compared to compound **29.**

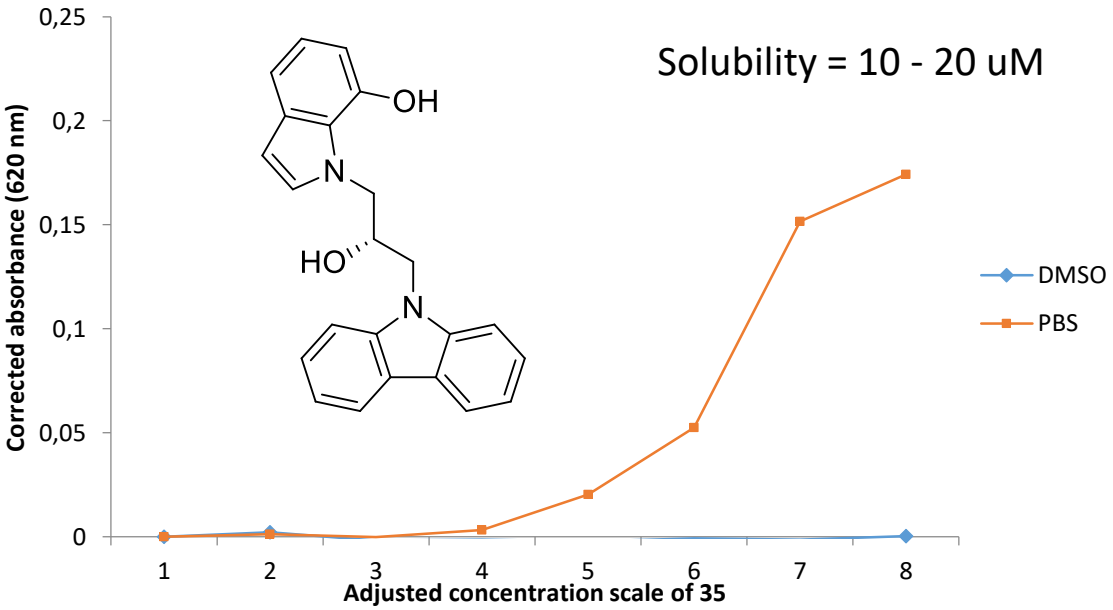


Figure 5.5: Corrected absorbance at 620 nm for **35** versus the adjusted concentration.

Compound **35** exhibited a solubility range of between 10 and 20 µM. This was due to the deviation of the curve from the baseline between the values of 3 (10.0 µM) and 4 (20.0 µM) on the x-axis. When compared to compound **31**, the solubility was increased due to the additional hydroxyl group present in the molecule, however, **35** was also termed moderately soluble.

5.3.2 β-haematin Inhibition Studies

As for the benzimidazole compounds in Chapter 4, the NP-40 detergent mediated β-haematin inhibition assay was modified for use on the carbazole indole compounds. Compounds **29**-**35** were tested using these conditions with the results shown in Table 5.1 alongside the Bayesian statistics results from Chapter 2 for purposes of comparison.

Table 5.1: The β-haematin inhibition results of compounds **29**-**35** and the accompanying Bayesian prediction from Chapter 2 with active compounds shown in red.

Compound	β-haematin inhibition IC_{50} (µM)	Bayesian β-haematin inhibition score
29	Inactive	-7.4793
30	97.4 ± 10.6	-5.29619
31	102.6 ± 2.7	-5.29619
32	96.6 ± 8.1	-5.29619
33	22.5 ± 2.4	4.72502
34	18.0 ± 3.2	4.72502
35	23.6 ± 2.3	4.72502

Analysing the data from the table, it can be seen that compounds **29**-**32** (compounds with no or one hydroxyl) can be termed inactive or very poorly active using the 100 µM cut-off, with the values given for comparison. A clear trend can be seen with these compounds where the β-haematin inhibition activity increases with the number of hydroxyl groups from zero (inactive) to one (inactive or poorly active) to two hydroxyls (active). Comparing the Bayesian predictions and experimental data, the two correlate completely with all seven compounds

correctly predicted leading to 100% hit rate with this series of compounds. This was a 30-fold enrichment compared to random screening of the β-haematin inhibition Bayesian model containing 2 133 active compounds and 64 118 inactive compounds. This showed that the Bayesian model performs for at least two scaffolds and is a crucial technique for the discovery of these compounds and derivatives thereof.

5.3.3 Docking Studies

Attempts were made to use molecular docking to explain the trend in β-haematin inhibition by using the intermolecular interactions approach used on the benzimidazole compounds. Using the same methodology as the benzimidazole compounds, the adsorption energies are given in Table 5.2.

Table 5.2: The adsorption energies obtained from the molecular docking on both β-haematin crystal faces of compounds **29, 31, 32, 34** and **35.**

Compound	β-haematin inhibition IC_{50} (μM)	Adsorption energy (kcal/mol) (001)	Adsorption energy (kcal/mol) (011)
29	Inactive	-	-
31	96.6 ± 8.1	-72.84	-54.81
32	102.6 ± 2.7	-66.44	-60.87
34	18.0 ± 3.2	-56.58	-59.92
35	23.6 ± 2.3	-64.58	-59.13

Using again the literature intermolecular interaction distances for hydrogen bonding, which were taken to be optimally between 2.36 and 3.69 Å,[164] and π-π interactions distances which were optimally taken to be between 3.3 and 3.8 Å,[165] the molecular docking results could be rationalised. The racemic compounds (**30**, **33**) were not simulated in this study as they are just mixtures of the enantiomers already simulated (**31**, **32**, **34**, **35**). Adsorption energies could not be obtained for **29** as the compound does not stay inside the β-haematin pocket during simulated annealing. From investigating the pre-annealing docking structure of compound **29,** (Figure 5.6) it can be seen that there are no hydrogen bonding capabilities for this

compound nor any feasible π-π interactions (4.6 Å and 4.8 Å is much greater than the literature maximum value of 3.8 Å). This correlates with the experimental inactivity of this compound. Examining the docked structures for compounds **31** and **32** (Figure 5.7), one long hydrogen bond can be seen between the hydroxyl group and a propionate group of the β-haematin crystal structure with a bond distance of 3.59 Å. The π-π interaction distances are too long with a value shown of 4.0 Å.

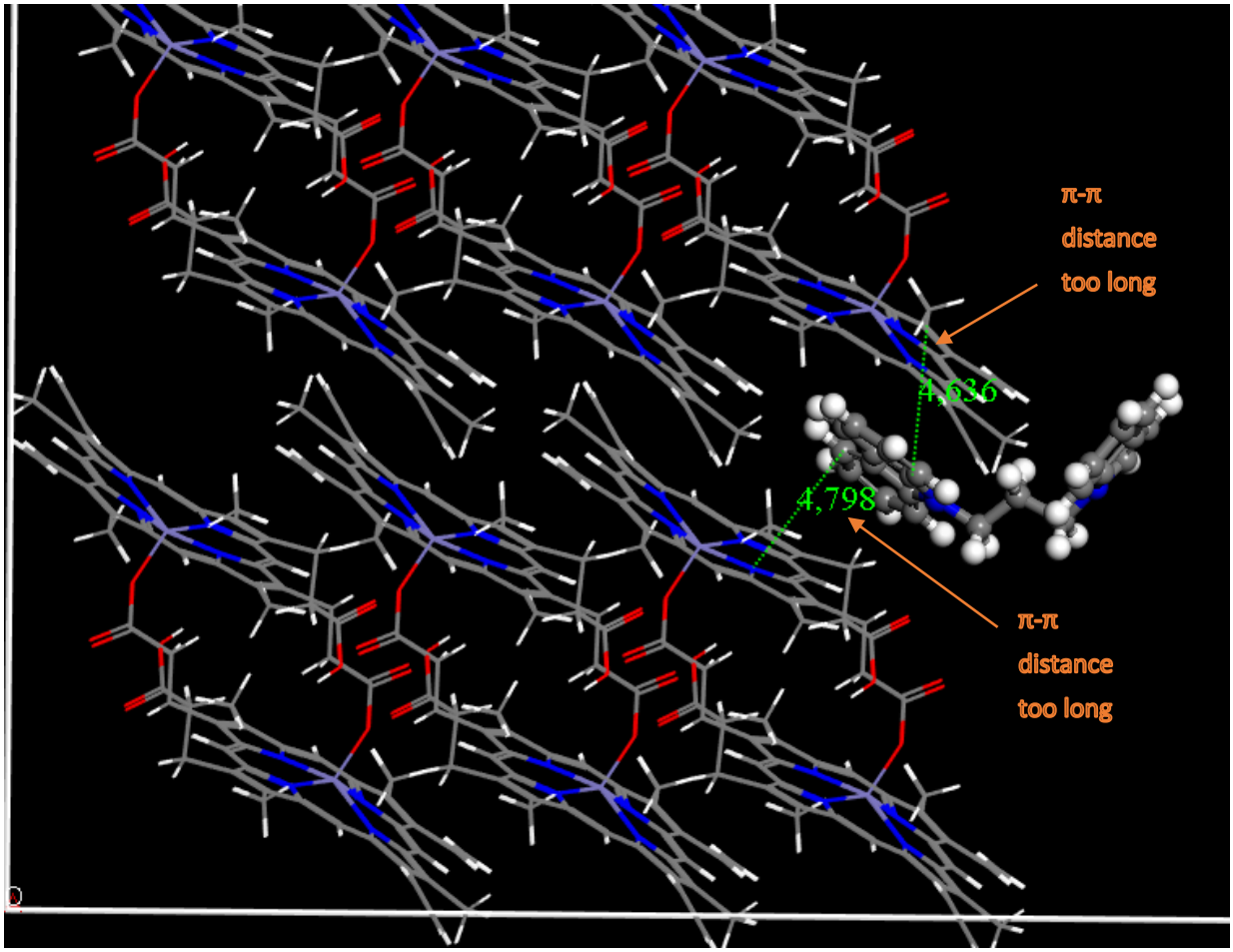


Figure 5.6: Docked structure of compound **29** on the (001) crystal face of the β-haematin crystal showing no intermolecular interactions with the β-haematin surface.

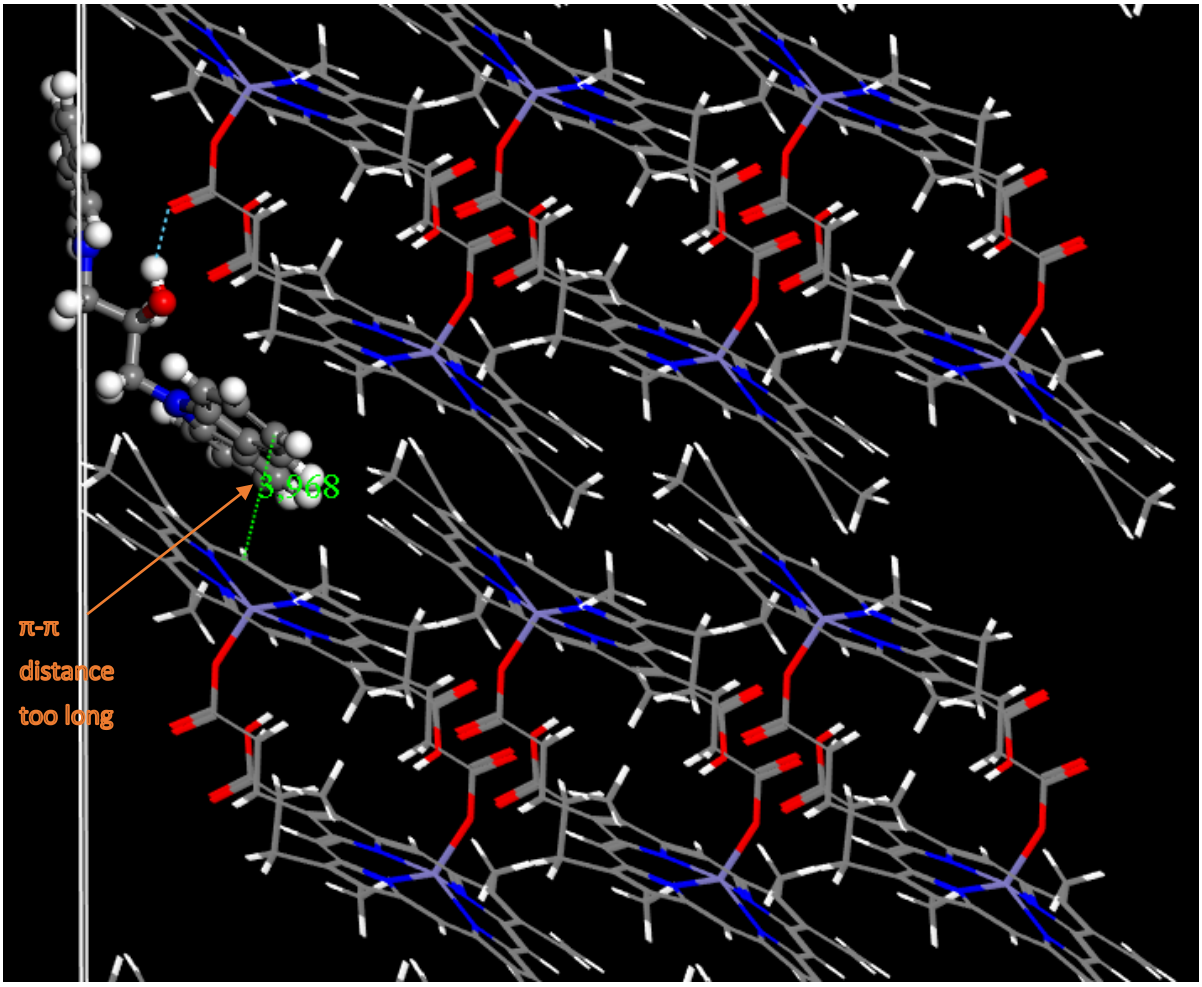


Figure 5.7: Docked structure of compound **31** on the (001) crystal face of the β-haematin crystal showing one hydrogen bond and no π-π intermolecular interactions.

Compounds **34** and **35** (Figure 5.8) show the same results as **31** and **32** except the hydroxyl of the indole functionality is involved in the hydrogen bond rather than alkyl hydroxyl. The short hydrogen bond (2.38 Å) between this functionality and the propionate group of the β-haematin crystal is shorter than compounds **31** and **32**, likely leading to a stronger bond (3.59 Å vs. 2.38 Å). This probably explains the increased experimental β-haematin inhibition activity.

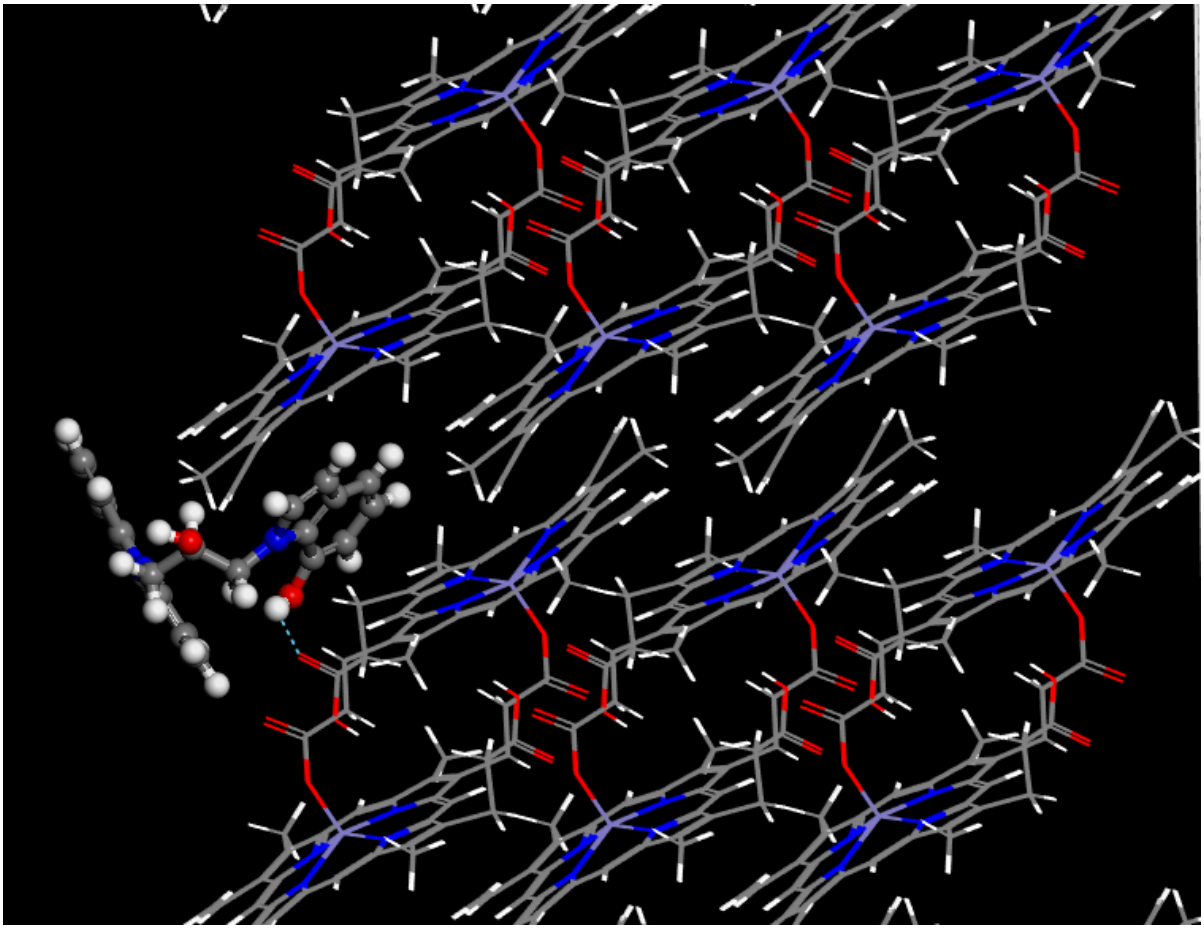


Figure 5.8: Docked structure of compound **35** on the (001) crystal face of the β-haematin crystal showing one hydrogen bond and no π-π intermolecular interactions.

It should also be noted that there are no significant intermolecular interactions with these compounds and the (011) β-haematin crystal face. An example of this is shown in Figure 5.9 where the π-π distances of 4.4 and 4.1 Å are much greater than the maximum value from literature of 3.8 Å. Therefore the (001) β-haematin crystal face was deemed the most important crystal face for these compounds.

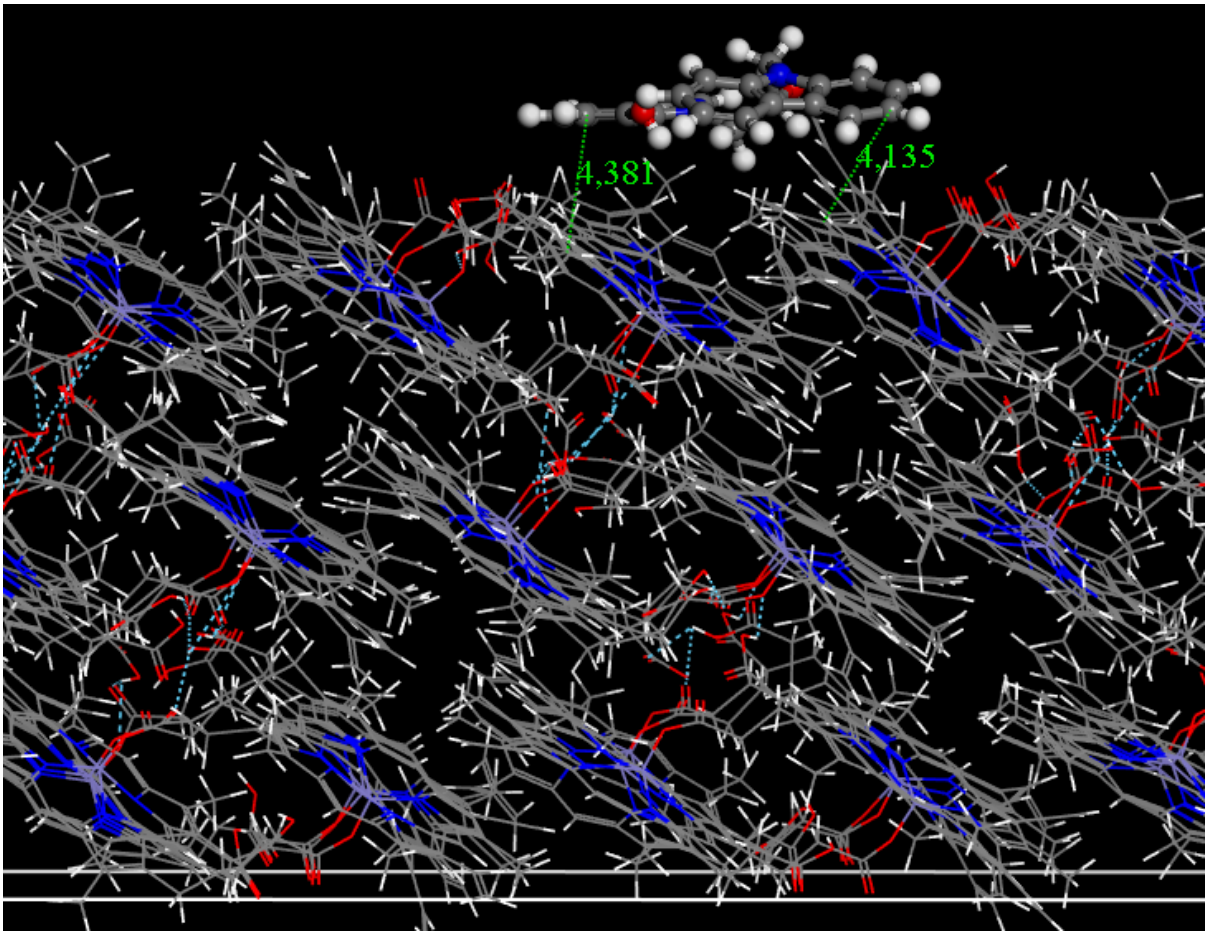


Figure 5.9: Docked structure of compound **35** on the (011) crystal face of the β-haematin crystal showing no intermolecular interactions.

There was no significant correlation found between the energy of interaction with the β-haematin crystal faces and the experimental β-haematin inhibition data due to insufficient data points.

5.3.4 Malaria Parasite Growth Inhibition Studies

Following the success of the β-haematin inhibition study and the molecular docking, the carbazole indole compounds were subjected to malaria parasite growth inhibition studies. These compounds were sent to collaborators and subjected to the same assay conditions as described in Chapter 4. The results of this study are shown in Table 5.3 with the relevant Bayesian statistics for comparison purposes.

All compounds in this series were shown to inhibit the parasite growth of the NF54 strain of the malaria parasite using a 20 µM cut-off, compared to the one inactive compound predicted by the Bayesian model. This led to a hit rate of 86%, but when compared to random screening

of the compounds in the malaria parasite growth inhibition Bayesian model containing 817 active and 41 729 inactive compounds, a 44-fold enrichment was achieved. However, the original Bayesian model contained a 2 µM cut-off so when compared to this all compounds are deemed to be inactive.

Table 5.3: The malaria parasite growth inhibition IC_{50} results of compounds **29**-**35** using the NF54 strain of *P. falciparum* and the accompanying Bayesian prediction from Chapter 2 with predicted active compounds shown in red.

Compound	NF54 malaria parasite growth inhibition IC_{50} (µM)	Malaria parasite growth inhibition score
29	14.15 ± 4.74	-10.3333
30	8.03 ± 2.94	-8.19434
31	6.82 ± 1.21	-8.19434
32	5.75 ± 0.66	-8.19434
33	8.86 ± 0.10	-5.38157
34	3.47 ± 0.10	-5.38157
35	7.64 ± 0.54	-5.38157

Compound **29** (no hydroxyl groups) showed the lowest malaria parasite growth inhibition activity of the series, while compounds **30**-**32** (one hydroxyl group) showed similar activities to compounds **33**-**35** (two hydroxyl groups). Of the different subgroups (**30**-**32** and **33**-**35**), the racemic compounds showed the lowest activity while **34** was found to be the most active compounds of the series.

Since all compounds showed some malaria parasite growth inhibition activity against the NF54 strain, it was prudent to analyse this series against the chloroquine resistant strain of *P. falciparum*, Dd2. The results of this are shown in Table 5.4 together with the NF54 results and resistance index. Compounds **29** (no hydroxyl groups) and **30**-**32** (one hydroxyl group) saw an increase in malaria parasite growth inhibition activity in the Dd2 strain of *P. falciparum*. The opposite trend was seen in compounds **33**-**35** (two hydroxyl groups). This

is clear when examining the resistance index where compounds **29**-**32** showed a resistance index of less than one compared to compounds **33**-**35** which showed a resistance index of greater than one. None of the compounds exhibited significant cross-resistance with chloroquine.

Table 5.4: The Dd2 malaria parasite growth inhibition IC_{50} results of compounds **29**-**35** and the accompanying NF54 malaria parasite growth inhibition IC_{50} results.

Compound	Dd2 malaria parasite growth inhibition IC_{50} (μM)	NF54 malaria parasite growth inhibition IC_{50} (μM)	Resistance Index (RI)
29	5.42 ± 1.01	14.15 ± 4.74	0.38
30	2.84 ± 0.36	8.03 ± 2.94	0.35
31	4.32 ± 0.36	6.82 ± 1.21	0.63
32	2.76 ± 0.10	5.75 ± 0.66	0.48
33	11.29 ± 0.90	8.86 ± 0.10	1.27
34	8.04 ± 0.50	3.47 ± 0.10	2.32
35	20.60 ± 7.73	7.64 ± 0.54	2.69

The cytotoxicity of compounds **29**-**35** was determined against the CHO cell line, with the results shown in Table 5.5. All compounds were shown to be selective against *P. falciparum* with the selectivity index (SI) ranging from 5.46 to 75.43 in the chloroquine resistant strain and 14.73 to 47.79 in the chloroquine sensitive strain.

In general, the malaria parasite growth inhibition activities results are poor in comparison to the original hit compound which showed low nanomolar values. This indicated that either the switch from the quinoline ring system to the indole ring system or the removal of the positive charge lowered activity substantially.

Table 5.5: The cytotoxicity results from the CHO cell line of compounds **29**-**35** with the selectivity indexes to the Dd2 and NF54 strain of *P. falciparum*.

Compound	CHO cytotoxicity (μM)	Dd2 malaria parasite growth inhibition IC_{50} (μM)	Selectivity Index (SI) Dd2	NF54 malaria parasite growth inhibition IC_{50} (μM)	Selectivity Index (SI) NF54
29	> 300.00	5.42 ± 1.01	> 55.35	14.15 ± 4.74	> 21.20
30	202.78 ± 97.14	2.84 ± 0.36	71.40	8.03 ± 2.94	25.25
31	218.15 ± 13.37	4.32 ± 0.36	50.50	6.82 ± 1.21	31.99
32	208.19 ± 7.34	2.76 ± 0.10	75.43	5.75 ± 0.66	36.21
33	183.21 ± 15.24	11.29 ± 0.90	16.23	8.86 ± 0.10	20.68
34	165.82 ± 35.41	8.04 ± 0.50	20.62	3.47 ± 0.10	47.79
35	112.53 ± 20.62	20.60 ± 7.73	5.46	7.64 ± 0.54	14.73

5.3.5 Cellular Haem Fractionation Assay

Following the biological analysis of the carbazole indole series, compounds **32** and **34** were sent to collaborators from the Egan research group for cellular haem fractionation analysis. The results for compound **32** are shown in Figure 5.10. Analysis showed that the percentage haemozoin remained constant throughout the study but the haemozoin iron content showed a large decrease. By contrast there was little to no change in the percentage free haem or haem iron content. This strongly indicated that this compound acts against a target that results in the inhibition of the uptake of haemoglobin into the food vacuole. This behaviour has previously been reported with other compounds such as atovaquone and hexahydroquinoline.[168,173]

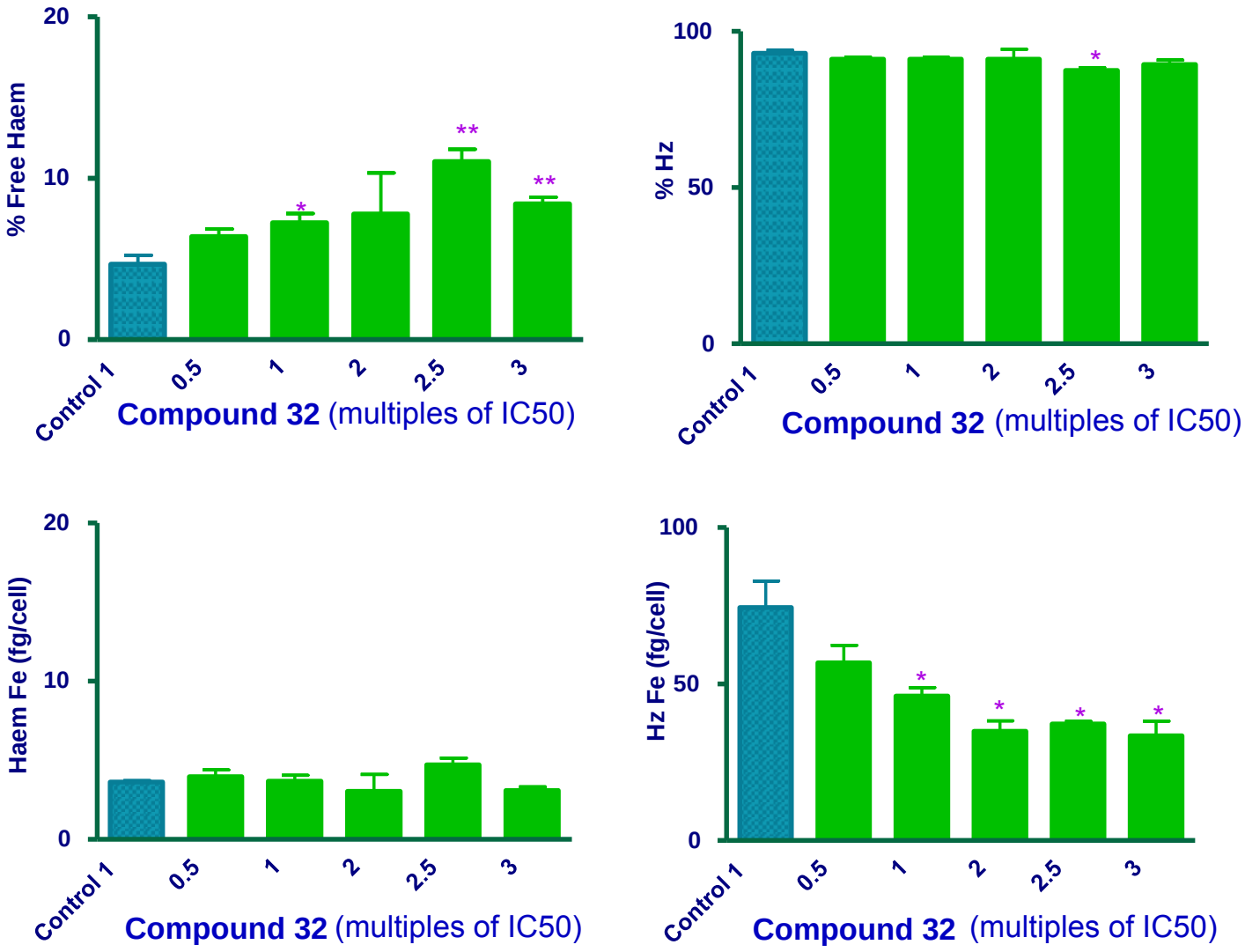


Figure 5.10: The percentage and iron content values of haem, and haemozoin after incubation with compound **32**.

Surprisingly, the same phenomenon occurred with compound **34** as shown in Figure 5.11. The percentage free haem increased, but the iron content of free haem in the cell remained constant as did the percentage haemozoin while the iron content of haemozoin in the cell decreased to about half the control value. The β-haematin inhibition IC_{50} of this compound was found to be 18.0 ± 3.2 µM, therefore it can be concluded that this series of compounds is probably not able to enter food vacuole or acts on a different target before reaching

concentrations adequate to inhibit haemozoin formation. Haemozoin inhibition is evidently not the mode of action of these compounds.

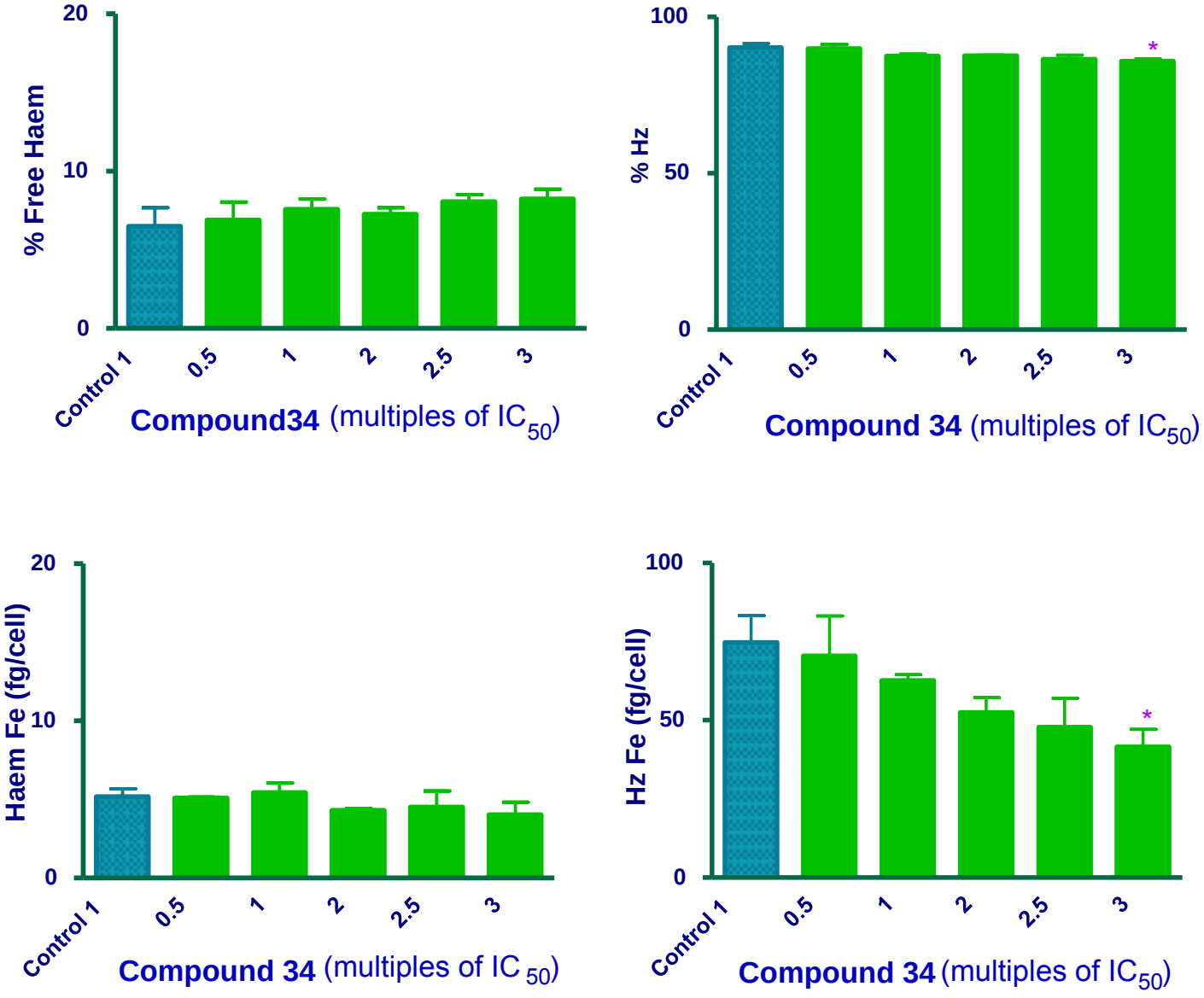


Figure 5.11: The percentage and iron content values of haem and haemozoin after incubation with compound **34**.

This shows how important this type of analysis is, as even though a compound is able to inhibit β-haematin formation, this might not be its ultimate mode of action.

5.4 Summary and Conclusions

In this study all the synthesized carbazole indole compounds were tested using the turbidimetric solubility assay. This revealed that an increase in hydroxyl groups increased the water solubility of the compounds. However, the compounds were only found to be moderately soluble with a maximum concentration range of 10 to 20 µM.

All synthesized carbazole indole compounds were also tested using the NP-40 detergent mediated β-haematin inhibition assay using the same procedure as used for the benzimidazoles. The series showed promising β-haematin inhibitory results with all compounds deemed active. Comparing this to the Bayesian predictions, a 100% hit rate was achieved proving yet again the utility of the method. When this was compared to random screening, a 30-fold enrichment was achieved. SAR analysis of this series revealed that an increase in the number of hydroxyl groups in the molecule increased the β-haematin inhibition activity.

Docking studies were also carried out on this series of compounds with the results showing that compound **29**, which was experimentally inactive in the β-haematin inhibition study, had no intermolecular interaction with both the fastest and second fastest growing β-haematin crystal surfaces. In contrast all the other compounds in the series contained one hydrogen bonding interaction with the (001) β-haematin crystal surface. This hydrogen bond was shorter for compounds **34** and **35**. It was also found that all compounds in the series did not interact with the (011) β-haematin crystal surface and thus the (001) β-haematin crystal surfaces was deemed the most important for this series of compounds.

P. falciparum growth inhibition studies were carried out on all compounds using the NF54 and Dd2 strains. Using a cut-off of 20 µM, all compounds were shown to inhibit the parasite growth of the NF54 strain. This led to an 86% hit rate compared to the Bayesian results with one compound incorrectly predicted. This was found to be a 44-fold enrichment compared to random screening using this model. However, using a cut-off of 2 µM all compounds were found to be inactive. Analysing the Dd2 results, compounds **29**-**32** were found to be more active when compared to the NF54 strain results with a resistance index of less than one. The opposite was found to be true for compounds **33**-**35.**

Haem fractionation results showed that the haemozoin iron content showed a large decrease while there was no change in the free haem levels or haem iron content. This indicated that these compounds have the effect of inhibiting the uptake of haemoglobin into the food vacuole, a phenomenon already seen in literature with other compounds. They are not haemozoin inhibitors and have an unknown mechanism of action.

Chapter 6: Conclusions and Future Work

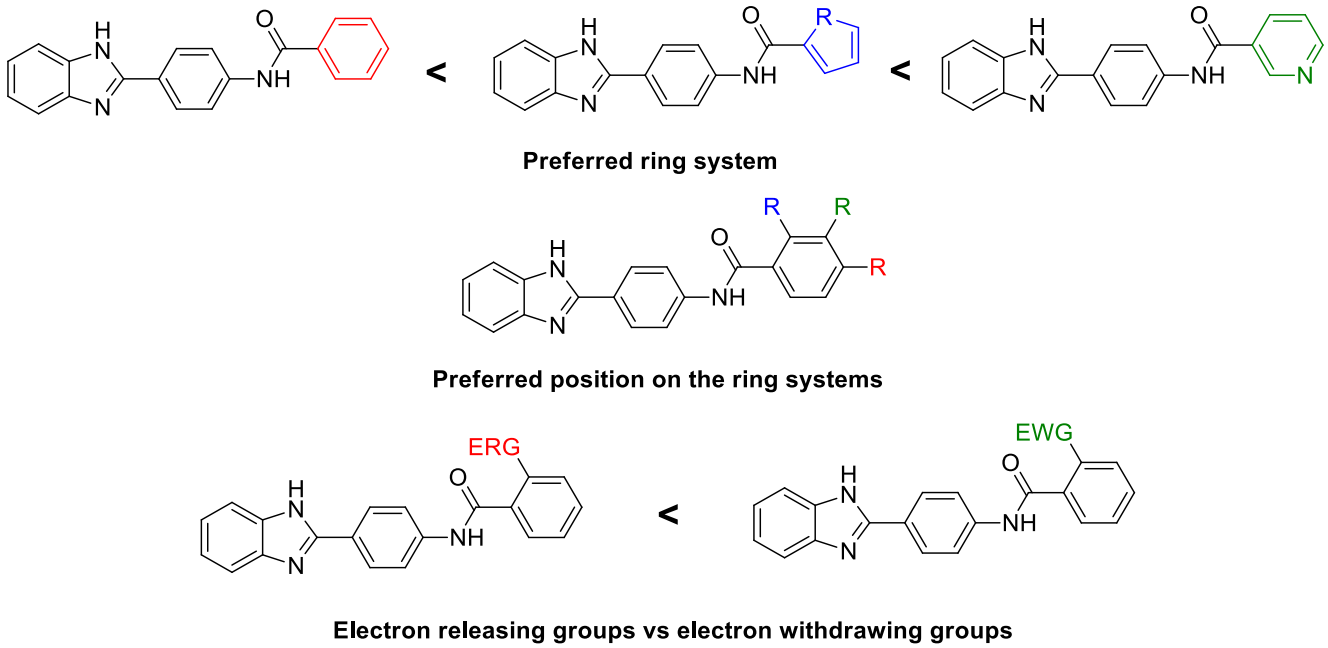
6.1 Conclusions

A drastic increase in malaria parasite resistance has caused decreased efficacy of well-known antimalarials such as chloroquine, mefloquine, amodiaquine and many others. There have also been reports of ACT resistance, even though efforts to minimise this are being implemented by the WHO. These challenges pose serious health risks globally thus resulting in a dire need for new, effective antimalarials comprising of novel scaffolds. To this end, a HTS at Vanderbilt University identified 171 compounds which were shown to be β-haematin inhibitors as well as inhibit greater than 90% parasitemia in *P. falciparum*. These 171 compounds were sorted by scaffold, of which the benzimidazole and carbazole scaffolds were selected for this project. One of the key factors for this choice was the two hit compounds, one for each scaffold, which appeared in the top ten most active compounds in the Vanderbilt HTS.

In this project, both the benzimidazole and carbazole scaffolds were further studied to discover SARs and probe possible mechanisms of action for the potential drug-like compounds using the hit compounds discussed as starting points for analysis.

Discovery Studio was used in conjunction with Pipeline Pilot and previously established Bayesian statistical models to enumerate and filter an *in silico* library consisting of benzimidazole derivatives, based on the hit compound, according to firstly β-haematin inhibition activity, followed by malaria parasite growth inhibition activity and finally aqueous solubility using the Sigma-Aldrich catalogue for purchasable reagents. After vigorous screening and compound elimination, 19 compounds were selected for a 2-step synthesis via a condensation reaction using PPA at 220 °C, followed by an acylation reaction using an optimised acid chloride methodology (2.0 eq. acid chloride, 1.0 eq. **1**, - 40 °C). A smaller series of seven benzimidazole compounds was also subjected to the same Bayesian models but using predetermined structures based on fragments from the original hit benzimidazole compound as well as a diacylated compound, **20**. These fragment compounds and **20** were either purchased or synthesized using similar acylation methodologies.

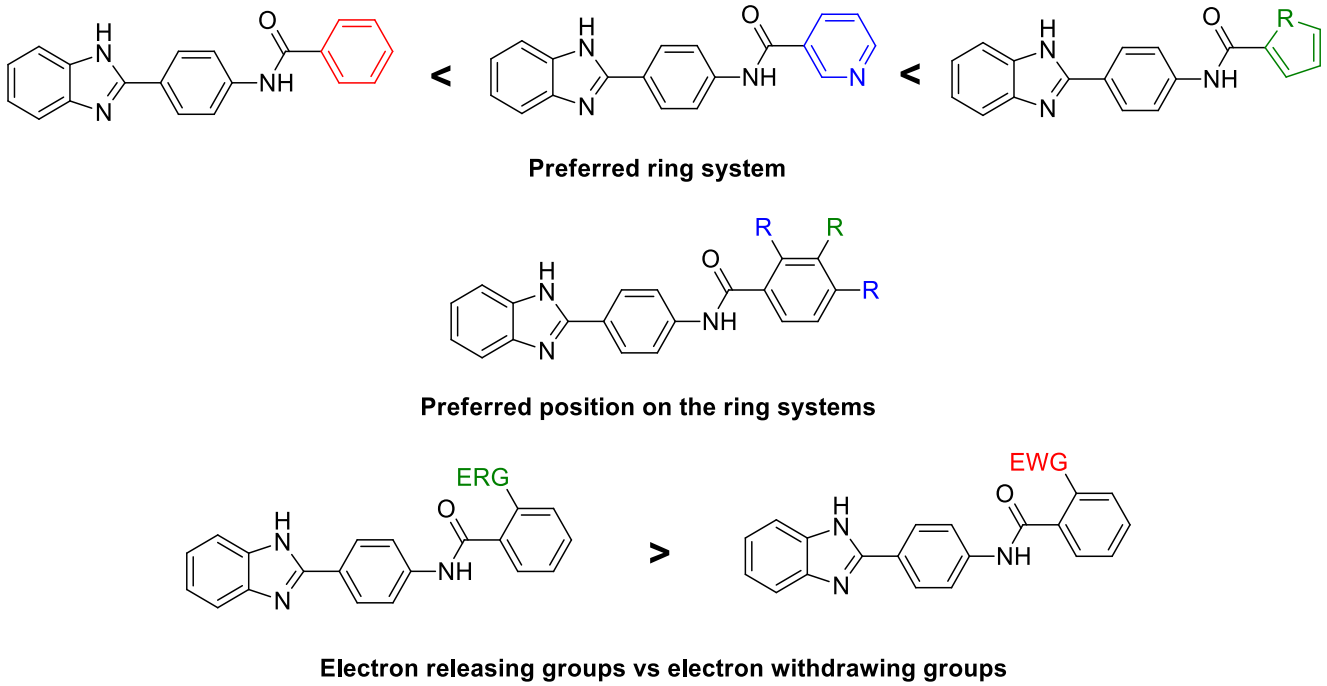
The benzimidazole compounds showed clear SAR trends for β-haematin inhibition and are summarised in Scheme 6.1. The overall preferred system for β-haematin inhibition was observed to be a pyridyl ring system with a *meta*-EWG substituent. *In silico* molecular docking of the 19 benzimidazole compounds on both the (001) and (011) β-haematin crystal faces revealed the (011) face to be the dominant face for adsorption according to a correlation with β-haematin inhibition activity and adsorption energies. This prompted an investigation into the intermolecular interaction between these compounds and both the β-haematin crystal faces which demonstrated that compounds must have a minimum of three π-π interactions between the compound and the crystal faces to be a β-haematin inhibitor.



Scheme 6.1: A summary of the SAR trends found for the benzimidazole series for β-haematin inhibition with good/ preferred groups in green, moderate groups in blue and poor groups in red.

Clear SAR trends observed from malaria parasite growth inhibition IC_{50} values are shown in Scheme 6.2. The ideal benzimidazole compound observed consisted of a 5-membered heteroaromatic ring system. *meta*-ERG substituents were found to be favoured on the 6-membered aromatic ring systems which is contradictory to the SAR trend found for β-haematin inhibition and thus a balance between these two trends must be found for a compound to have potent activity. None of the compounds exhibited cross-resistance when tested against the chloroquine resistant strain with disubstituted compounds found to be

more active against the chloroquine resistant strain of *P. falciparum*. All benzimidazole fragment compounds were found to be inactive when tested in the parasite growth inhibition assay. Thus, it was concluded that all functional groups, not just the benzimidazole functionality, are required for *in vitro* activity.



Scheme 6.2: The SAR trends found for the malaria parasite growth inhibition IC_{50} of the benzimidazole series with preferred groups in green, moderate groups in blue and poor groups in red.

Cellular haem fractionation studies confirmed that the mode of action of the benzimidazole compounds is haemozoin inhibition with an increase in free haem and a decrease in haemozoin shown in the parasite. A multiple correlation study further showed for an increase in malaria parasite growth inhibition activity the β-haematin activity, number of hydrogen bond donors and molecular depth must also increase.

Seven carbazole indole compounds were selected with varying hydrogen bonding ability to analyse the influence of this on β-haematin inhibition as well as parasite growth inhibition activities. These activities were first predicted using the same Bayesian statistical models in Material Studio as with the benzimidazole compounds. The carbazole indole compounds were synthesized either via a 2-step or 4-step synthetic pathway, the former involving a NaH mediated S_N2 alkylation using carbazole and an epichlorohydrin stereoisomer or

1,3-dibromopropane, followed by another NaH mediated S_N2 reaction to open the epoxide or eliminate bromine and furnish the desired compounds. The 4-step synthetic sequence was similar, but in addition to the previous steps mentioned also involved benzyl protection of 7-hydroxyindole prior to epoxide opening followed by deprotection using hydrogen gas and Pd-C. Solubility studies were performed on these compounds using the turbidimetric solubility assay. This was possible due to their increased aqueous solubility compared to the benzimidazole compounds. The best carbazole indole compound was found to be moderately aqueous soluble with a solubility range of 10-20 μM.

β-haematin inhibition analysis on the carbazole indole compounds demonstrated that there was a definite trend in activity from no hydroxyl groups (inactive), one hydroxyl group (poor activity) and two hydroxyl groups (active). Docking analysis of the carbazole indole compounds demonstrated that these compounds only interact via hydrogen bonding with the (001) β-haematin crystal face. It was shown that hydrogen bonding interactions form the basis of β-haematin inhibition activity in the carbazole indole compounds as there were no π-π interactions shown between these compounds and both the β-haematin crystal surfaces. A similar trend was shown in the malaria parasite growth inhibition study, where the compound without any hydroxyl group was shown to be inactive while the other six compounds showed moderate activities. These *in vitro* activities where substantially weaker than the original hit compound and thus either the quinoline ring system, the positive charge or a combination of the two are essential for activity. *In vitro* testing on the chloroquine resistant strain showed no cross resistance with the compounds with two hydroxyl groups showing a decrease in activity compared to the chloroquine sensitive strain. Cellular haem fractionation studies demonstrated that haemozoin inhibition is not the mode of action of the carbazole indole compounds but rather follows similar behaviour to atovaquone.

Analysis of all the predicted Bayesian activities for both scaffolds demonstrated a 4-fold to 44-fold enrichment compared to random screening within the respective Bayesian model. This showed the benefits of using these Bayesian models to screen compounds before attempting to synthesize the respective compounds.

In this project, two novel non-quinoline containing scaffolds were investigated revealing their fairly simple synthetic pathways. The scaffold characteristics relating to β-haematin inhibition, adsorption onto the β-haematin crystal surfaces, parasite growth inhibition

activities, free haem levels and ultimately the mechanism of action was revealed. This data provides insight into these scaffolds so that future drug design can be guided.

6.2 Future Work

The benzimidazole hit compound found during the HTS was still found to be the most active of the series even after investigation into other derivatives. However, now that SAR for this scaffold has been established further attempts to increase the activity of this scaffold can be made. The interplay between the SAR for β-haematin inhibition and the SAR for malaria parasite growth inhibition should be investigated. Mainly, the question of whether a 5-membered ring system or a pyridyl ring system is best should be further investigated. Further derivatives guided by the new SAR input in the Bayesian models should be attempted especially looking at substituted 5-membered rings in light of the *in vitro* SAR demonstrated in this project. One of the main challenges with this scaffold was the overall low aqueous solubility of the compounds. Attempts to improve this solubility with the addition of hydrophilic functional groups will enable simpler synthesis as well as better ADME properties of this scaffold.

Only seven compounds were synthesized for the carbazole indole compounds. Due to this small series, SAR trends other than hydrogen bonding were not possible, therefore it would be beneficial to expand on this study by exploring various other derivatives of this scaffold. As with the benzimidazole scaffold, the Bayesian models, which are now further enriched with the data from this project, would assist in the choice of derivatives. Optimisation or a new synthetic route are needed to improve the enantiomeric excess of these compounds in order to make these compounds viable for further testing. This preliminary study also demonstrated that these compounds do not act via haemozoin inhibition but rather have a different mode of action. It would be interesting to investigate the original hit compound, even though it contained an unfavourable positive charge, to show if the initial carbazole quinoline scaffold also showed the same results. Furthermore, it has been shown that this carbazole scaffold is an ideal candidate for a drug discovery programme due to its potent activity shown in the HTS as well as reasonable aqueous solubility demonstrated in this project which will more likely be significantly improved with a charged compound. For these reasons an in-depth study of this scaffold is needed.

Due to the complex nature of *P. falciparum* resistance and treatment, all resources and data available must be used to aid in the design and discovery of novel antimalarial compounds to ensure that, hopefully in the near future the scourge of malaria can finally be eradicated.

Chapter 7: Experimental

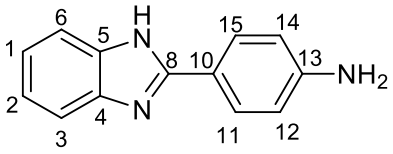
7.1 General Synthetic Methods

All reagents were purchased from commercial sources (Sigma-Aldrich, Kimix, and Protea chemicals). All solvents except anhydrous pyridine were distilled before use. Tetrahydrofuran in particular was distilled under nitrogen using sodium wire and benzophenone. Calcium chloride was used as a drying reagent in drying tubes. Thin layer chromatography was performed using aluminium-backed silica gel 60 F254 plates, which were purchased from Merck and visualised using a UV lamp, anisaldehyde spray (a mix in a 1:1 (v/v) ratio of 10% H_2SO_4 in ethanol solution and 5% anisaldehyde in ethanol solution) or a ninhydrin spray (300 mg Ninhydrin in 97 mL absolute ethanol and 3 mL of glacial acetic acid). Column chromatography was carried out using silica gel 60 (mesh 68 – 280 µm) purchased from Fluka and a Biotage Isolera One flash chromatography system. Nuclear magnetic resonance experiments were recorded using a Bruker or Varian Unity 300, 400 or 600 MHz instrument. Chemical shifts (δ) were recorded relative to residual DMSO-d_6 (δ 2.50 in ^{1}H NMR and δ 39.52 in ^{13}C NMR), chloroform-*d* (δ 7.26 in ^{1}H NMR and δ 77.16 in ^{13}C NMR), methanol-d_4 (δ 4.87 in ^{1}H NMR and δ 49.00 in ^{13}C NMR) or acetone-d_6 (δ 2.05 in ^{1}H NMR and δ 206.26 in ^{13}C NMR). All chemical shifts were reported in ppm and all *J* values were reported in Hz. Melting points were obtained using a Reichert-Jung Thermovar hot stage microscope. HPLC experiments were carried out using an Agilent Technologies 1220 Infinity LC with a reverse phase C18 column using a solvent system of double distilled deionised Millipore© Direct-Q water and HPLC grade acetonitrile or a Daicel Chiracel OD (250 x 4.6 mm) column using HPLC grade hexane and HPLC grade isopropanol as the solvent system. High resolution mass spectrometry was performed on the Waters Synapt G2 instrument. All mass spectra were recorded using the electrospray positive (ES^+) technique and the sample was introduced via an ESI probe injected into a stream of acetonitrile.

7.2 Benzimidazole Compounds

7.2.1 Precursor

4-(1*H*-Benzo[*d*]imidazol-2yl)aniline (1)



1

4-Aminobenzoic acid (5.00 g, 36.46 mmols), *o*-phenylenediamine (3.94 g, 36.46 mmols) and PPA (50.0 g, 10 eq.) were added to a 50 mL reaction vessel with a drying tube and heated at 220 °C, with stirring for 5 hr. The reaction progress was monitored by TLC using a 6:94 mixture of MeOH:DCM. The resulting mixture was washed with 200 mL hot water into a 10% $NaHCO_3$ solution (200 mL) and basified yielding a brown solid. This solid was washed with water (3 x 400 mL) and crystallized from methanol and water to yield light brown needle crystals of (**1**) (4.68 g, 61% yield).

ATR-FTIR ν_{max} /cm^{-1} 3360, 3441 (NH_2); M.p. 240 - 241 °C (lit. 240 - 241 °C)[174]; R_f (MeOH/DCM 6:94) 0.49; δ_H (DMSO-d_6, 400 MHz) 5.54 (s, 2H, NH_2), 6.68 (d, *J* = 8.6 Hz, 2H, H-12, H-14), 7.11 (m, 2H, H-1, H-2), 7.49 (m, 2H, H-3, H-6), 7.86 (d, *J* = 8.6 Hz, 2H, H-11, H-15); δ_C (DMSO-d_6, 100.6 MHz) 113.5 (C_H), 113.5 (C_H), 114.2 (C_{quat}), 117.3 (C_{quat}), 121.2 (C_H), 127.7 (C_H), 150.5 (C_{quat}), 152.6 (C_{quat}); HRMS-ES^+ Calculated: 210.1031 $[M+H]^+$ for $C_{13}H_{12}N_3$, Observed: 210.1031; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.1%.

7.2.2 Benzimidazole Derivatives

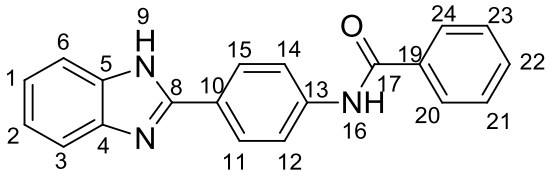
7.2.2.1 Procedure for the preparation of acid chlorides

Using a typical scale of 2.0 mmols, the appropriate aryl carboxylic acid (2.0 mmols) was added to a small round-bottomed flask, which was flushed with nitrogen. Thionyl chloride (3 mL, 41.4 mmols) was added after which the reaction mixture was refluxed, with stirring, at 80 °C and left overnight to ensure complete conversion. The reaction mixture was cooled and excess thionyl chloride was completely removed under reduced pressure. The crude acid chloride was then used as a reagent for subsequent reactions.

7.2.2.2 Synthetic procedure for preparation of *N*-(4-1*H*-benzo[*d*]imidazol-2-yl)phenyl)benzamides from the appropriate acid chlorides

Using a typical scale of 1.0 mmol, **1** (1.0 mmol) was dissolved in anhydrous pyridine (2 mL) in a small round-bottomed flask and the mixture was cooled to -40 °C after which the aryl acid chloride (2.0 mmols; purchased or prepared as above) in dry THF or DMF (2 mL) was added dropwise over 20 min with vigorous stirring. After 2-6 hr the reaction mixture was warmed to room temperature and the solvent reduced under reduced pressure. Without using a work-up the crude product was purified by column chromatography directly using MeOH/DCM mixtures (1:99 to 2:8).

***N*-4-(1*H*-Benzo[*d*]imidazol-2-yl)phenyl)benzamide (2)**

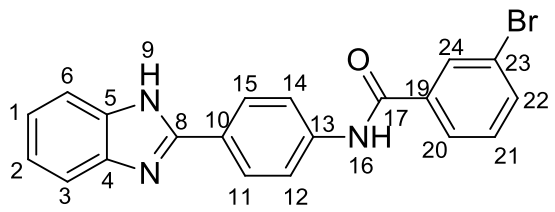


2

Benzoyl chloride (197 mg, 1.4 mmols) and **1** (146 mg, 0.7 mmols) afforded a off-white solid**.** The resulting solid was recrystallized from methanol to afford cream crystals of (**2**) (132 mg, 60% yield).

ATR-FTIR v_{max} /cm^{-1} 1641 (C=O amide); M.p. 340 - 341 °C (lit. M.p 335 °C)[175]; R_f (MeOH/DCM 5:95) 0.38; δ_H (DMSO-d_6, 400 MHz) 7.28 (m, 2H, H-1, H-2), 7.53-7.63 (m, 3H, H-21, H-22, H-23), 7.65 (m, 2H, H-3, H-6), 7.98-8.02 (m, 2H, H-20, H-24), 8.03 (d, *J* = 8.9 Hz, 2H, H-12, H-14), 8.22 (d, *J* = 8.9 Hz, 2H, H-11, H-15), 10.54 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 114.6 (C_H), 120.3 (C_H), 122.8 (C_H), 123.3 (C_{quat}), 127.4 (C_H), 127.7 (C_H), 128.4 (C_H), 131.7 (C_{quat}), 134.6 (C_{quat}), 137.4 (C_{quat}), 141.4 (C_{quat}), 150.5 (C_{quat}), 165.8 (CO); HRMS-ES^+ Calculated: 314.1293 $[M+H]^+$ for $C_{20}H_{16}N_3O$, Observed: 314.1293; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.2%.

***N*-(4-(1*H*-benzo[*d*]imidazole-2-yl)phenyl)-3-bromobenzamide (3)**

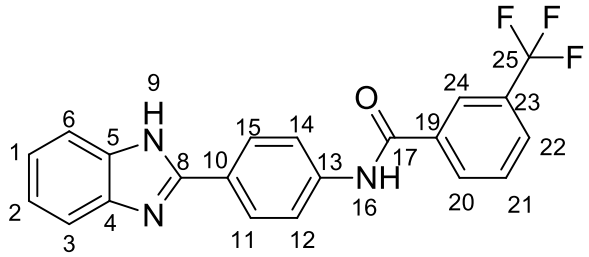


3

3-Bromobenzoic acid (804 mg, 4.0 mmols) and **1** (418 mg, 2.0 mmols) afforded a brown solid**.** The resulting solid was recrystallized from methanol to afford light-brown crystals of (**3**) (668 mg, 85% yield).

ATR-FTIR ν_{max} /cm^{-1} 3248 (NH), 1652 (C=O amide); M.p. 324 - 325 °C; R_f (MeOH/DCM 5:95) 0.39; δ_H (DMSO-d_6, 400 MHz) 7.20 (m, 2H, H-1, H-2), 7.63 – 7.79 (m, 3H, H-3, H-6, H-21), 7.81 (d, *J* = 7.8 Hz, 1H, H-22), 7.98 (m, 3H, H-12, H-14, H-20), 8.19 (m, 3H, H-11, H-15, H-24), 10.54 (s, 1 H, H-16), 12.80 (s, 1 H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 111.1 (C_H), 118.6 (C_H), 120.4 (C_H), 121.5 (C_H), 121.7 (C_H), 122.3 (C_{quat}), 125.6 (C_{quat}), 126.9 (C_H), 126.9 (C_H), 130.3 (C_H), 130.6 (C_H), 134.4 (C_H), 135.0 (C_{quat}), 136.9 (C_{quat}), 140.3 (C_{quat}), 143.9 (C_{quat}), 151.1 (C_{quat}), 164.1 (CO); HRMS-ES$^+$ Calculated: 392.0398 $[M+H]^+$ for $C_{20}H_{15}Br^{79}N_3O$, Observed: 392.0399; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.8%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-3-(trifluoromethyl)benzamide (4)**



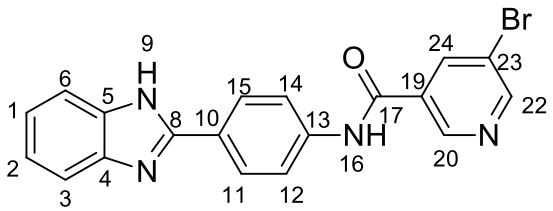
4

3-(Trifluoromethyl)benzoic acid (266 mg, 1.4 mmols) and **1** (146 mg, 0.7 mmols) afforded a white solid**.** The resulting solid was recrystallized from ethanol to afford white crystals of (**4**) (119 mg, 45% yield).

ATR-FTIR ν_{max} /cm^{-1} 3267 (NH), 1655 (C=O amide); M.p. 302 – 303 °C; R_f (MeOH/DCM 5:95) 0.30; δ_H (DMSO-d_6, 400 MHz) 7.16-7.23 (m, 2H, H-1, H-2), 7.52 (d, *J* = 7.4 Hz, 1H, H-6), 7.65 (d,

J = 7.1 Hz, 1H, H-3), 7.81 (t, *J* = 7.8 Hz, 1H, H-21), 7.96-8.00 (m, 3H, H-20, H-12, H-14), 8.20 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 8.30 (d, *J* = 7.8 Hz, 1H, H-22), 8.34 (s, 1H, H-24), 10.66 (s, 1H, H-16), 12.81 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 111.1 (C_H), 118.6 (C_H), 120.5 (C_H), 121.5 (C_H), 122.3 (C_H), 124.2 (C_H), 125.7 (C_{quat}), 125.3 (C_{quat}), 126.9 (C_H), 128.2 (C_H), 129.2 (q, *J* = 127.7 Hz, C-25), 129.7 (C_H), 131.8 (C_H), 135.0 (C_{quat}), 135.6 (C_{quat}), 140.2 (C_{quat}), 143.9 (C_{quat}), 151.0 (C_{quat}), 164.2 (CO); HRMS-ES$^+$ Calculated: 382.1167 $[M+H]^+$ for $C_{21}H_{15}N_3OF_3$, Observed: 382.1166; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.6%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-5-bromonicotinamide (5)**

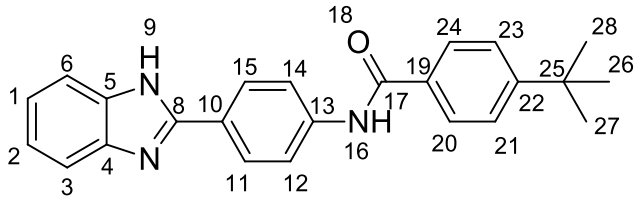


5

5-Bromopyridine-3-carboxylic acid (283 mg, 1.4 mmols) and **1** (146 mg, 0.7 mmols) afforded a white solid**.** The resulting solid was recrystallized from ethanol to afford white crystals of (**5**) (211 mg, 77% yield).

ATR-FTIR ν_{max} /cm^{-1} 3306 (NH), 1653 (C=O amide); M.p. 325-326 °C; R_f (MeOH/DCM 1:9) 0.82; δ_H (DMSO-d_6, 400 MHz) 7.20 (m, 2H, H-1, H-2), 7.58 (m, 2H, H-3, H-6), 7.95 (d, *J* = 8.8 Hz, 2H, H-12, H-14), 8.19 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 8.58 (t, *J* = 2.1 Hz, 1H, H-24), 8.92 (d, *J* = 2.2 Hz , 1H, H-20), 9.10 (d, *J* = 1.9 Hz, 1H, H-22), 10.69 (s, 1H, H-16), 12.83 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 111.3 (C_H), 118.7 (C_H), 120.0 (C_{quat}), 120.3 (C_H), 121.8 (C_H), 121.8 (C_H), 125.8 (C_{quat}), 127.0 (C_H), 132.0 (C_{quat}), 135.0 (C_{quat}), 137.7 (C_H), 140.0 (C_{quat}), 143.8 (C_{quat}), 147.4 (C_H), 151.0 (C_{quat}), 152.8 (C_H), 162.7 (CO); HRMS-ES$^+$ Calculated: 393.0351 $[M+H]^+$ for $C_{19}H_{14}BrN_4O$, Observed: 393.0354; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.4%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-4-(tert-butyl)benzamide (6)**

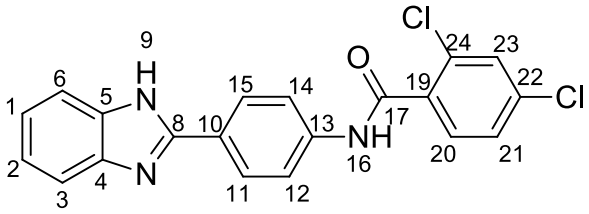


6

4-*tert*-Butylbenzoic acid (357 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a white solid**.** The resulting solid was recrystallized from ethanol to afford white crystals of (**6**) (297 mg, 87% yield).

ATR-FTIR v_{max} /cm^{-1} 1652 (C=O amide); M.p. 317-318 °C; R_f (MeOH/DCM 7:93) 0.92; δ_H (DMSO-d_6, 400 MHz) 1.34 (s, 9H, H-26, H-27, H-28), 7.52 (m, 2H, H-1, H-2), 7.57 (d, *J* = 8.6 Hz, 2H, H-12, H-14), 7.81 (m, 2H, H-3, H-6), 7.95 (d, *J* = 8.6 Hz, 2H, H-11, H-15), 8.13 (d, *J* = 8.9 Hz, 2H, H-21, H-23), 8.36 (d, *J* = 8.9 Hz, 2H, H-20, H-24), 10.67 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 30.9 (C_{H3}), 34.7 (C_{quat}), 113.8 (C_H), 118.0 (C_{quat}), 120.2 (C_H), 125.2 (C_H), 125.4 (C_H), 127.7 (C_H), 128.8 (C_H), 131.7 (C_{quat}), 132.3 (C_{quat}), 143.7 (C_{quat}), 148.7 (C_{quat}), 154.9 (C_{quat}), 166.0 (CO); HRMS-ES^+ Calculated: 370.1919 $[M+H]^+$ for $C_{24}H_{24}N_3O$, Observed: 370.1914; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.7%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-2,4-dichlorobenzamide (7)**



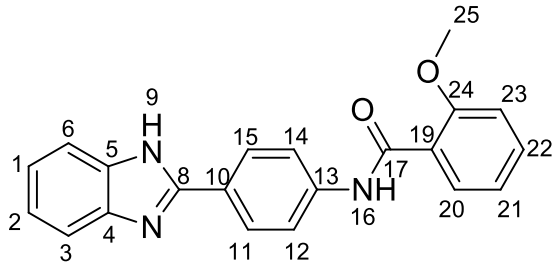
7

2,4-Dichlorobenzoic acid (267 mg, 1.4 mmols) and **1** (146 mg, 0.7 mmols) afforded a brown solid**.** The resulting solid was recrystallized from methanol to afford light- brown crystals of (**7**) (186 mg, 70% yield).

ATR-FTIR v_{max} /cm^{-1} 1657 (C=O amide); M.p. 273-275 °C (lit. M.p 242 °C)[175]; R_f (MeOH/DCM 1:9) 0.50; δ_H (DMSO-d_6, 400 MHz) 7.16-7.24 (m, 2H, H-1, H-2), 7.50-7.60 (m, 2H, H-3, H-20), 7.62-7.70 (m, 2H, H-6, H-21), 7.78 (m, 1H, H-23), 7.89 (d, *J* = 8.6 Hz, 2H, H-12, H-14), 8.18 (d, *J*

= 8.6 Hz, 2H, H-11, H-15), 10.75 (s, 1H, H-16), 12.81 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 111.1 (C_H), 118.6 (C_H), 119.6 (C_H), 121.5 (C_H), 122.3 (C_H), 125.7 (C_{quat}), 127.1 (C_H), 127.4 (C_H), 129.2 (C_H), 130.3 (C_H), 131.2 (C_{quat}), 135.0 (C_{quat}), 135.5 (C_{quat}), 140.0 (C_{quat}), 143.8 (C_{quat}), 143.8 (C_{quat}), 151.0 (C_{quat}), 164.1 (CO); HRMS-ES$^+$ Calculated: 382.0514 $[M+H]^+$ for $C_{20}H_{14}Cl_2N_3O$, Observed: 382.0508; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.6%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-2-methoxybenzamide (8)**

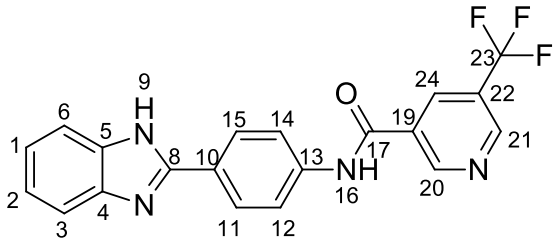


8

2-Methoxybenzoic acid (304 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol and water to afford dark brown crystals of (**8**) (280 mg, 82% yield).

ATR-FTIR ν_{max} /cm^{-1} 3331 (NH), 1661 (C=O amide); M.p. 133-134 °C; R_f (MeOH/DCM 1:9) 0.93; δ_H (DMSO-d_6, 400 MHz) 3.93 (s, 3H, H-25), 7.09 (t, *J* = 7.4 Hz, 1H, H-21), 7.18-7.23 (m, 3H, H-1, H-2, H-20), 7.52 (m, 1H, H-22), 7.59 (m, 2H, H-3, H-6), 7.69 (m, 1H, H-23), 7.93 (d, *J* = 8.4 Hz, 2H, H-11, H-15), 8.16 (d, *J* = 8.4 Hz, 2H, H-12, H-14), 10.32 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 55.6 (C_{H3}), 112.0 (C_H), 114.8 (C_H), 119.7 (C_H), 120.5 (C_H), 121.9 (C_H), 124.7 (C_{quat}), 125.0 (C_{quat}), 127.0 (C_H), 129.7 (C_H), 132.2 (C_H), 139.2 (C_{quat}), 140.5 (C_{quat}), 151.1 (C_{quat}), 156.5 (C_{quat}), 164.6 (CO); HRMS-ES$^+$ Calculated: 344.1399 $[M+H]^+$ for $C_{21}H_{18}N_3O_2$, Observed: 344.1393; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.3%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-3-(trifluoromethyl)benzamide (9)**

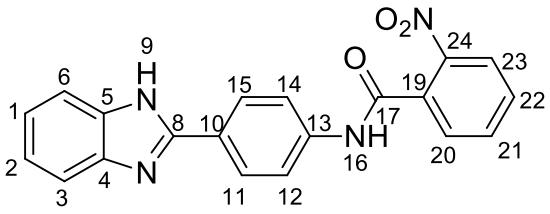


9

5-(Trifluoromethyl)pyridine-3-carboxylic acid (191 mg, 1.0 mmol) and **1** (105 mg, 0.5 mmols) afforded a white solid**.** The resulting solid was recrystallized from ethanol to afford white crystals of (**9**) (116 mg, 68% yield).

ATR-FTIR v_{max} /cm^{-1} 1661 (C=O amide); M.p. 318-320 °C; R_f (MeOH/DCM 1:9) 0.11; δ_H (DMSO-d_6, 400 MHz) 7.20 (m, 2H, H-1, H-2), 7.59 (m, 2H, H-3, H-6), 7.97 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 8.21 (d, *J* = 8.8 Hz, 2H, H-12, H-14), 8.71 (m, 1H, H-24), 9.20 (m, 1H, H-20), 9.41 (m, 1H, H-21), 10.81 (s, 1H, H-16); δ_C (DMSO- d_6, 100.6 MHz) 115.1 (C_H), 120.4 (C_H), 122.0 (C_H), 124.9 (q, *J* = 127.8 Hz, C_{quat}), 125.8 (C_{quat}), 127.0 (C_H), 130.5 (C_{quat}), 132.7 (C_{quat}), 132.7 (C_H), 139.9 (C_{quat}), 148.6 (C_H), 148.6 (C_{quat}), 150.9 (C_{quat}), 152.6 (C_H), 162.7 (CO); HRMS-ES^+ Calculated: 383.1120 $[M+H]^+$ for $C_{20}H_{14}F_3N_4O$, Observed: 383.1117; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 95.0%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-2-nitrobenzamide (10)**



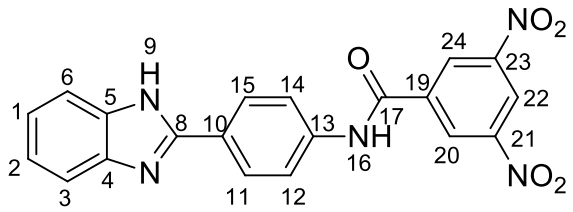
10

2-Nitrobenzoyl chloride (223 mg, 1.2 mmols) and **1** (126 mg, 0.6 mmols) afforded a yellow solid**.** The resulting solid was recrystallized from methanol to afford yellow crystals of (**10**) (145 mg, 67% yield).

ATR-FTIR v_{max} /cm^{-1} 1652 (C=O amide), 1527 (N-O stretch); M.p. 308-309 °C; R_f (MeOH/DCM 5:95) 0.36; δ_H (DMSO-d_6, 400 MHz) 7.16-7.24 (m, 2H, H-1, H-2), 7.48-7.57 (m, 1H, H-6), 7.61-

7.69 (m, 1H, H-3), 7.74-7.93 (m, 5H, H-11, H-15, H-20, H-21, H-22), 8.14-8.22 (m, 3H, H-12, H-14, H-23), 10.80 (s, 1H, H-16), 12.73 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 110.9 (C_H), 118.5 (C_H), 119.6 (C_H), 121.3 (C_H), 122.1 (C_H), 124.0 (C_H), 125.6 (C_{quat}), 127.0 (C_H), 129.1 (C_H), 130.8 (C_H), 132.4 (C_{quat}), 133.9 (C_H), 134.9 (C_{quat}), 140.0 (C_{quat}), 143.8 (C_{quat}), 146.3 (C_{quat}), 150.9 (C_{quat}), 164.1 (CO); HRMS-ES$^+$ Calculated: 359.1144 $[M+H]^+$ for $C_{20}H_{15}N_4O_3$, Observed: 359.1151; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.8%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-3,5-dinitrobenzamide (11)**

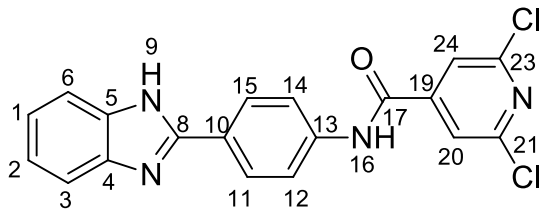


11

3,5-Dinitrobenzoyl chloride (461 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a yellow solid**.** The resulting solid was recrystallized from methanol to afford yellow crystals of (**11**) (355 mg, 89% yield).

ATR-FTIR ν_{max} /cm^{-1} 3304 (NH), 1661 (C=O amide), 1531 (N-O stretch); M.p. 338-339 °C; R_f (MeOH/DCM 1:9) 0.68; δ_H (DMSO-d_6, 400 MHz) 7.48 (m, 2H, H-1, H-2), 7.78 (m, 2H, H-3, H-6), 8.12 (d, J = 8.4 Hz, 2H, H-12, H-14), 8.37 (d, J = 8.4 Hz, 2H, H-11, H-15), 9.02 (s, 1H, H-22), 9.20 (s, 2H, H-20, H-24), 11.25 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 113.7 (C_H), 119.9 (C_{quat}), 120.5 (C_H), 120.9 (C_H), 124.6 (C_H), 127.8 (C_H), 128.2 (C_H), 133.2 (C_{quat}), 136.8 (C_{quat}), 141.8 (C_{quat}), 147.8 (C_{quat}), 148.6 (C_{quat}), 161.5 (CO); HRMS-ES$^+$ Calculated: 404.0995 $[M+H]^+$ for $C_{20}H_{14}N_5O_5$, Observed: 404.1004; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.2%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-2,6-dichloroisonicotinamide (12)**

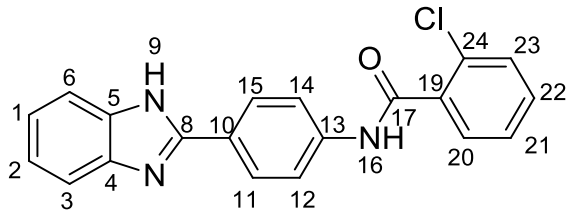


12

2,6-Dichloropyridine-4-carboxylic acid (384 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a yellow solid**.** The resulting solid was recrystallized from methanol and water to afford pale yellow crystals of (**12**) (264 mg, 69% yield).

ATR-FTIR ν_{max} /cm^{-1} 1664 (C=O amide); M.p. 348-349 °C; R_f (MeOH/DCM 1:9, 1% Et_3N) 0.32; δ_H (DMSO-d_6, 400 MHz) 7.20 (m, 2H, H-1, H-2), 7.59 (m, 2H, H-3, H-6), 7.93 (d, *J* = 8.8 Hz, 2H, H-12, H-14), 8.03 (s, 2H, H-20, H-24), 8.20 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 10.80 (s, 1H, H-16), 12.80 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 111.1 (C_H), 111.4 (C_H), 118.3 (C_H), 120.4 (C_H), 121.8 (C_H), 121.8 (C_H), 126.2 (C_{quat}), 127.0 (C_H), 139.5 (C_{quat}), 144.1 (C_{quat}), 144.1 (C_{quat}), 148.0 (C_{quat}), 149.8 (C_{quat}), 150.9 (C_{quat}), 161.3 (CO); HRMS-ES$^+$ Calculated: 383.0466 $[M+H]^+$ for $C_{19}H_{13}Cl_2N_4O$, Observed: 383.0468; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.7%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-2-chlorobenzamide (13)**

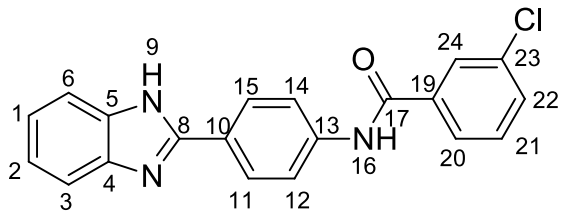


13

2-Chlorobenzoic acid (313 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol to afford brown crystals of (**13**) (184 mg, 53% yield).

ATR-FTIR ν_{max} /cm^{-1} 1657 (C=O amide); M.p. 275-276 °C (lit. M.p 311 °C)[175]; R_f (MeOH/DCM 6:94) 0.95; δ_H (DMSO-d_6, 400 MHz) 7.46-7.51 (m, 1H, H-21), 7.53–7.57 (m, 3H, H-1, H-2, H-22), 7.58-7.61 (m, 1H, H-20), 7.63-7.67 (m, 1H, H-23), 7.82 (m, 2H, H-3, H-6), 8.04 (d, *J* = 8.8 Hz, 2H, H-12, H-14), 8.44 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 11.03 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 113.8 (C_H), 118.0 (C_{quat}), 119.7 (C_H), 125.6 (C_H), 127.3 (C_H), 129.0 (C_H), 129.2 (C_H), 129.7 (C_H), 129.9 (C_{quat}), 131.4 (C_H), 131.9 (C_{quat}), 136.3 (C_{quat}), 143.2 (C_{quat}), 148.4 (C_{quat}), 165.5 (CO); HRMS-ES$^+$ Calculated: 348.0904 $[M+H]^+$ for $C_{20}H_{15}ClN_3O$, Observed: 348.0901; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 100%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-3-chlorobenzamide (14)**

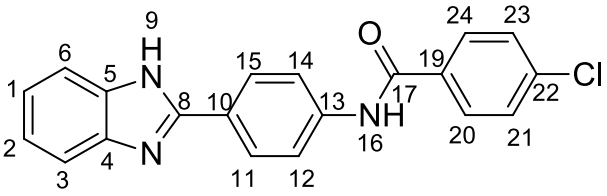


14

3-Chlorobenzoic acid (313 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a cream solid**.** The resulting solid was recrystallized from methanol to afford off-white crystals of (**14**) (319 mg, 92% yield).

ATR-FTIR v_{max} /cm^{-1} 1669 (C=O amide); M.p. 320-322 °C (lit. M.p 312 °C)[175]; R_f (MeOH/DCM 6:94) 0.95; δ_H (DMSO-d_6, 400 MHz) 7.50-7.55 (m, 2H, H-1, H-2), 7.58 (t, *J* = 7.9 Hz, 1H, H-21), 7.67 (m, 1H, H-22), 7.79-7.84 (m, 2H, H-3, H-6), 7.99 (dt, *J* = 1.4, 7.8 Hz, 1H, H-20), 8.07 (t, *J* = 1.8 Hz, 1H, H-24), 8.13 (d, *J* = 8.9 Hz, 2H, H-12, H-14), 8.42 (d, *J* = 8.9 Hz, 2H, H-11, H-15), 10.78 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 113.6 (C_H), 118.0 (C_{quat}), 120.3 (C_H), 125.3 (C_H), 126.4 (C_H), 127.3 (C_H), 128.7 (C_H), 130.1 (C_H), 131.4 (C_H), 132.0 (C_{quat}), 133.1 (C_{quat}), 136.1 (C_{quat}), 143.1 (C_{quat}), 148.4 (C_{quat}), 164.4 (CO); HRMS-ES^+ Calculated: 348.0904 $[M+H]^+$ for $C_{20}H_{15}ClN_3O$, Observed: 348.0899; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 95.9%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)-4-chlorobenzamide (15)**



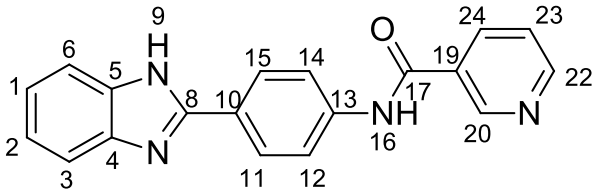
15

4-Chlorobenzoic acid (313 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol to afford brown crystals of (**15**) (260 mg, 75% yield).

ATR-FTIR v_{max} /cm^{-1} 1650 (C=O amide); M.p. 331-333 °C (lit. M.p 323 °C)[175]; R_f (MeOH/DCM 6:94) 0.94; δ_H (DMSO-d_6, 400 MHz) 7.49-7.54 (m, 2H, H-1, H-2), 7.61 (d, *J* = 8.6 Hz 2H, H-21,

H-23), 7.79-7.84 (m, 2H, H-3, H-6), 8.06 (d, *J* = 8.6 Hz, 2H, H-20, H-24), 8.12 (d, *J* = 8.9 Hz, 2H, H-12, H-14), 8.40 (d, *J* = 8.9 Hz, 2H, H-11, H-15), 10.70 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 113.5 (C_H), 118.1 (C_{quat}), 120.2 (C_H), 125.1 (C_H), 128.1 (C_H), 128.5 (C_H), 129.4 (C_H), 132.2 (C_{quat}), 132.9 (C_{quat}), 136.5 (C_{quat}), 143.1 (C_{quat}), 148.5 (C_{quat}), 164.7 (CO); HRMS-ES$^+$ Calculated: 348.0904 $[M+H]^+$ for $C_{20}H_{15}ClN_3O$, Observed: 348.0898; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 100%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)nicotinamide (16)**

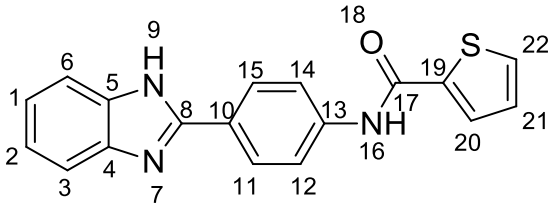


16

Nicotinic acid (246 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol and water to afford light brown crystals of (**16**)
(211 mg, 67% yield).

ATR-FTIR ν_{max} /cm^{-1} 3276 (NH), 1648 (C=O amide); M.p. 332-333 °C; R_f (MeOH/DCM 15:85) 0.17; δ_H (DMSO-d_6, 400 MHz) 7.16-7.24 (m, 2H, H-1, H-2), 7.50-7.72 (m, 3H, H-3, H-6, H-23), 7.97 (m, 2H, H-12, H-14), 8.12-8.24 (m, 2H, H-11, H-15), 8.30-8.36 (m, 1H, H-24), 8.76-8.81 (m, 1H, H-22), 9.14-9.19 (m, 1H, H-20), 10.64 (s, 1H, H-16), 12.81 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 115.1 (C_H), 120.2 (C_H), 121.6 (C_H), 123.2 (C_H), 125.6 (C_{quat}), 126.8 (C_H), 130.3 (C_{quat}), 135.1 (C_H), 140.0 (C_{quat}), 142.9 (C_{quat}), 148.4 (C_H), 150.9 (C_{quat}), 151.9 (C_H), 164.0 (CO); HRMS-ES$^+$ Calculated: 315.1246 $[M+H]^+$ for $C_{19}H_{15}N_4O$, Observed: 315.1249; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 97.7%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)thiophene-2-carboxamide (17)**

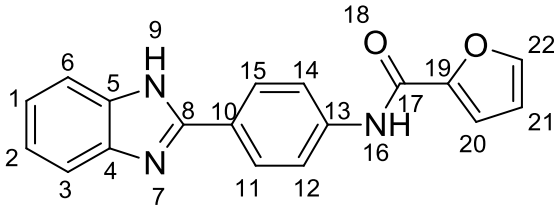


17

2-Thiophenecarboxylic acid (256 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol and water to afford light brown crystals of (**17**) (158 mg, 49% yield).

ATR-FTIR v_{max} /cm^{-1} 1650 (C=O amide); M.p. 344-346 °C; R_f (MeOH/DCM 1:9) 0.95; δ_H (DMSO-d_6, 400 MHz) 7.24 (t, *J* = 4.3 Hz, 1H, H-21), 7.42-7.47 (m, 2H, H-1, H-2), 7.73-7.79 (m, 2H, H-3, H-6), 7.87 (d, *J* = 5.0 Hz, 1H, H-22), 8.05-8.10 (d, *J* = 8.7 Hz, 2H, H-12, H-14), 8.17 (d, *J* = 3.8 Hz, 1H, H-20), 8.29-8.34 (d, *J* = 8.7 Hz, 2H, H-11, H-15), 10.62 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 113.8 (C_H), 119.5 (C_{quat}), 120.1 (C_H), 124.4 (C_H), 127.8 (C_H), 128.2 (C_H), 129.7 (C_H), 132.1 (C_H), 133.7 (C_{quat}), 139.2 (C_{quat}), 142.4 (C_{quat}), 149.1 (C_{quat}), 160.1 (CO); HRMS-ES$^+$ Calculated: 320.0858 $[M+H]^+$ for $C_{18}H_{14}SN_3O$, Observed: 320.0862; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 97.3%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)furan-2-carboxamide (18)**



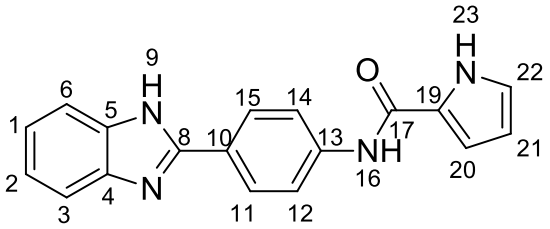
18

2-Furoic acid (224 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol to afford a light brown solid of (**18**) (232 mg, 75% yield).

ATR-FTIR v_{max} /cm^{-1} 1650 (C=O amide); M.p. 252-253 °C (lit. M.p 301 °C)[175]; R_f (MeOH/DCM 1:9) 0.86; δ_H (DMSO-d_6, 400 MHz) 6.72 (dd, *J* = 1.7, 3.5 Hz, 1H, H-21), 7.49 (d, *J* = 3.5 Hz, 1H,

H-22), 7.51-7.56 (m, 2H, H-1, H-2), 7.79-7.84 (m, 2H, H-3, H-6), 7.95 (m, 1H, H-20), 8.12 (d, *J* = 8.9 Hz, 2H, H-12, H-14), 8.37 (d, *J* = 8.9 Hz, 2H, H-11, H-15), 10.60 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 112.0 (C_H), 113.6 (C_H), 115.5 (C_H), 117.9 (C_{quat}), 120.1 (C_H), 125.3 (C_H), 128.7 (C_H), 132.1 (C_{quat}), 142.9 (C_{quat}), 145.9 (C_H), 146.9 (C_{quat}), 148.5 (C_{quat}), 156.3 (CO); HRMS-ES$^+$ Calculated: 304.1086 $[M+H]^+$ for $C_{18}H_{14}N_3O_2$, Observed: 304.1072; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 95.1%.

***N*-(4-(1-Benzo[*d*]imidazole-2-yl)phenyl)1*H*-pyrrole-2-carboxamide (19)**



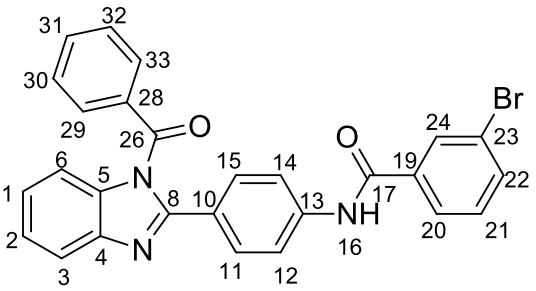
19

Pyrrole-2-carboxylic acid (222 mg, 2.0 mmols) and **1** (209 mg, 1.0 mmol) afforded a brown solid**.** The resulting solid was recrystallized from methanol and water to afford a light brown solid of (**19**) (162 mg, 54% yield).

ATR-FTIR ν_{max} /cm^{-1} 1654 (C=O amide); M.p. 292-293 °C; R_f (MeOH/DCM 1:9) 0.61; δ_H (DMSO-d_6, 400 MHz) 6.21 (m, 1H, H-21), 7.03 (m, 1H, H-20), 7.17 (m, 1H, H-22), 7.49-7.54 (m, 2H, H-1, H-2), 7.77-7.83 (m, 2H, H-3, H-6), 8.08-8.13 (d, *J* = 8.9 Hz, 2H, H-12, H-14), 8.25-8.30 (d, *J* = 8.9 Hz, 2H, H-11, H-15), 10.28 (s, 1H, H-16), 11.84 (s, 1H, H-9); δ_C (DMSO-d_6, 100.6 MHz) 109.1 (C_H), 112.5 (C_H), 113.8 (C_H), 117.6 (C_{quat}), 119.7 (C_H), 123.3 (C_H), 125.2 (C_H), 125.6 (C_{quat}), 128.6 (C_H), 132.7 (C_{quat}), 143.8 (C_{quat}), 149.1 (C_{quat}), 159.3 (CO); HRMS-ES$^+$ Calculated: 303.1246 $[M+H]^+$ for $C_{18}H_{15}N_4O$, Observed: 303.1233; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 95.0%.

7.2.3 Fragment Benzimidazole Derivatives

***N*-(4-(1-Benzoyl-1*H*-benzo[*d*]imidazole-2-yl)phenyl)-3-bromobenzamide (20)**

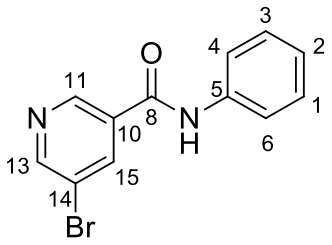


20

Compound **3** (300 mg, 0.76 mmols)**,** was added to a solution of anhydrous pyridine (2 mL) and THF (2 mL) in a 25 mL reaction vessel under nitrogen. Benzoyl chloride (0.8 mL, 6.9 mmols) was added and the solution left to stir for 18 hr at room temperature. The reaction mixture was quenched with MeOH, acidified with 1M HCl to pH 5 and the combined organic components extracted, first using MeOH: chloroform (1:3, 20 mL) and then chloroform (2x 20 mL). The organic extracts were dried over anhydrous Na_2SO_4 and the crude product purified using column chromatography with MeOH: DCM (1:99 to 7:93) mixtures. The resulting solid was further recrystallized from acetone and H_2O which afforded a light brown solid of (**20**) (146 mg, 39% yield)

ATR-FTIR ν_{max} /cm^{-1} 1657 (C=O amide); M.p. 222-223 °C; R_f (MeOH/DCM 1:9) 0.90; δ_H (DMSO-d_6, 400 MHz) 7.30-7.33 (m, 2H, H-1, H-2), 7.35-7.41 (m, 1H, H-3), 7.44-7.52 (m, 3H, H-6, H-12, H-14), 7.60-7.65 (m, 3H, H-11, H-15, H-21), 7.75-7.80 (m, 5H, H-29, H-30, H-31, H-32, H-33), 7.84 (dt, *J* = 0.9, 7.9 Hz, 1H, H-22), 7.94 (ddd, *J* = 1.0, 1.6, 7.8 Hz, 1H, H-20), 8.13 (t, *J* = 1.8 Hz, 1H, H-24), 10.45 (s, 1H, H-16); δ_C (DMSO-d_6, 100.6 MHz) 112.7 (C_H), 119.6 (C_H), 119.7 (C_H), 121.6 (C_{quat}), 124.2 (C_H), 124.4 (C_H), 125.3 (C_{quat}), 126.8 (C_H), 128.9 (C_H), 129.6 (C_H), 130.2 (C_H), 130.3 (C_H), 130.6 (C_H), 132.7 (C_{quat}), 134.3 (C_H), 134.4 (C_H), 134.6 (C_{quat}), 136.7 (C_{quat}), 140.2 (C_{quat}), 142.6 (C_{quat}), 153.2 (C_{quat}), 164.1 (CO), 168.9 (CO); HRMS-ES$^+$ Calculated: 496.0661 $[M+H]^+$ for $C_{27}H_{19}BrN_3O_2$, Observed: 496.0654; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 96.6%.

5-Bromo-*N*-phenylnicotinamide (21)



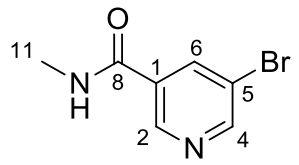
21

5-Bromopyridine-3-carboxylic acid (404 mg, 2.0 mmols) was added to a 25 mL reaction vessel and flushed with nitrogen. Thionyl chloride (2 mL) was added under dry conditions and the resulting mixture was stirred at 80 °C, overnight to ensure complete conversion to the aryl acid chloride. The reaction mixture was cooled and excess thionyl chloride was completely removed under pressure.

Aniline (102 mg, 1.1 mmols) was added to a 25 mL reaction flask, dissolved in anhydrous pyridine (2 mL) and cooled to -40 °C under inert conditions. A mixture of the previously formed 5-bromopyridine-3-carbonyl chloride (≈ 2 mmols) and THF (1.5 mL) was added dropwise over 20 min, with vigourous stirring and the reaction mixture slowly warmed to room temperature. After complete conversion, determined by TLC, the resulting mixture was quenched with MeOH and basifiied to pH 7 using anhydrous $NaHCO_3$. The combined organic components were extracted using DCM (3x 20 mL) and dried over Na_2SO_4. The crude product was then purified by column chromatography using MeOH/DCM (1:99 to 2:98) mixtures. The resulting compound was recrystallized from DCM to yield white crystals of (**21**) (287 mg, 94% yield).

ATR-FTIR v_{max} /cm^{-1} 3294 (NH), 1650 (C=O amide); M.p. 182 - 183 °C; R_f (MeOH/DCM 1:99) 0.21; δ_H (DMSO-d_6, 400 MHz) 7.11-7.16 (m, 1H, H-2), 7.35-7.40 (m, 2H, H-1, H-3), 7.75 (m, 2H, H-4, H-6), 8.54 (m, 1H, H-15), 8.90 (m, 1H, H-13), 9.07 (m, 1H, H-11), 10.47 (s, 1H, NH) ; δ_C (DMSO- d_6, 100.6 MHz) 119.9 (C_{quat}), 120.3 (C_H), 124.1 (C_H), 128.6 (C_H), 132.1 (C_{quat}), 137.6 (C_H), 138.5 (C_{quat}), 147.3 (C_H), 152.6 (C_H), 162.4 (CO); HRMS-ES^+ Calculated: 276.9977 $[M+H]^+$ for $C_{12}H_{10}BrN_2O$, Observed: 276.9974; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.5%.

5-Bromo-*N*-methylnicotinamide (22)



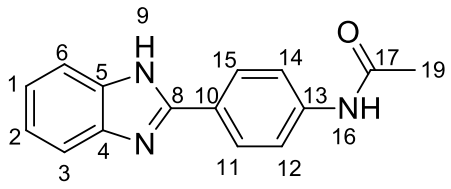
22

5-Bromopyridine-3 carboxylic acid (303 mg, 1.5 mmols) was added to a 25 mL reaction vessel and flushed with nitrogen. Thionyl chloride (4 mL) was added under dry conditions and the resulting mixture was stirred at 80 °C, overnight to ensure complete conversion to the aryl acid chloride. The reaction mixture was cooled and excess thionyl chloride was completely removed under pressure.

Methylamine.HCl (50 mg, 0.75 mmols) and freshly distilled, anhydrous triethylamine (1 mL) was added to a 25 mL reaction flask, which was added dropwise to the mixture of the previously formed 5-bromopyridine-3-carbonyl chloride in dry THF (2 mL). The mixture was allowed to stir vigourously at room temperature under nitrogen until complete conversion, via TLC monitoring, after which the mixture was acidified to pH 7 using 1 M HCl and the organic material extracted using MeOH:chloroform (1:1, 20 mL) followed by chloroform (2x 20 mL). The combined extracts were dried over anhydrous Na_2SO_4 and the crude product purified by column chromatography using MeOH/DCM (2:99 to 7:93) mixtures. The resulting solid was recrystallized using acetone to yield light yellow crystals of (**22**) (81 mg, 50% yield).

ATR-FTIR ν_{max} /cm^{-1} 3294 (NH), 1641 (C=O amide); M.p. 142 - 143 °C; R_f (MeOH/DCM 5:95) 0.17; δ_H (Acetone-d_6, 400 MHz) 2.93 (d, *J* = 4.7 Hz, 3H, H-11), 7.96 (s, 1H, NH), 8.35 (t, *J* = 2.1 Hz, 1H, H-6), 8.78 (d, *J* = 2.2 Hz, 1H, H-4), 8.98 (m, 1H, H-2); δ_C (Acetone-d_6, 100.6 MHz) 26.8 (CH_3), 121.1 (C_{quat}), 133.0 (C_{quat}), 138.2 (C_H), 147.7 (C_H), 153.5 (C_H), 165.0 (CO); HRMS-ES^+ Calculated: 214.9820 $[M+H]^+$ for $C_7H_8N_2OBr$, Observed: 214.9814; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 99.5%.

***N*-(4-(1*H*-Benzo[*d*]imidazole-2-yl)phenyl)acetamide (23)**



23

A solution of compound **1** (209 mg, 1.0 mmol), acetic anhydride (0.1 mL, 1.1 mmols) and toluene (3 mL) was refluxed at 115 °C for 2 hr and subsequently cooled. The precipitate was filtered and washed with hot ethyl acetate and toluene to give (**23**) (216 mg, 86% yield) as a white solid.

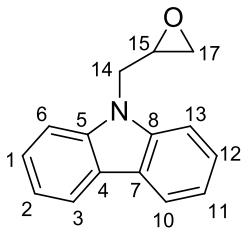
ATR-FTIR v_{max} /cm^{-1} 1673 (C=O amide); M.p. 302 - 303 °C (lit. 308 – 309 °C)[156]; R_f (MeOH/DCM 5:95) 0.30; δ_H (DMSO-d_6, 600 MHz) 2.09 (s, 3H, H-19), 7.18 (m, 2H, H-1, H-2), 7.57 (m, 2H, H-3, H-6), 7.75 (d, *J* = 8.8 Hz, 2H, H-12, H-14), 8.10 (d, *J* = 8.8 Hz, 2H, H-11, H-15), 10.16 (s, 1H, H-9); δ_C (DMSO- d_6, 150.9 MHz) 24.1 (CH_3), 114.9 (C_H), 118.9 (C_H), 121.9 (C_H), 124.5 (C_{quat}), 127.1 (C_H), 139.1 (C_{quat}), 140.8 (C_{quat}), 151.1 (C_{quat}), 168.6 (CO); HRMS-ES$^+$ Calculated: 252.1137 $[M+H]^+$ for $C_{15}H_{14}N_3O$, Observed: 252.1135; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (40:60), 98.6%.

7.3 Carbazole Derivatives

7.3.1 General Procedure for the Synthesis of Epoxide Precursors

To a solution of NaH (482 mg, 12.05 mmols) in anhydrous DMF (2 mL) at 0 °C was added a solution of carbazole (2.00 g, 11.97 mmols) in anhydrous DMF (3 mL) dropwise over 20 min. The reaction mixture was warmed to room temperature and stirred for 1 hr to allow for complete deprotonation. Thereafter, epichlorohydrin (1.00 mL, 13.17 mmols) was added to the reaction mixture dropwise and stirred for 16 hr. The reaction mixture was quenched with a saturated NH_4Cl solution and the combined organic components were extracted using EtOAc (2x 20 mL). The combined extracts were washed with H_2O (9x 10 mL) and brine (10 mL) and then dried over Na_2SO_4. The crude product was purified by column chromatography using EtOAc/ hexane (0:100 to 1:9) mixtures and the resulting solid was recrystallized from ethanol to yield colourless crystals as needles.

9-(Oxiran-2-ylmethyl)-9*H*-carbazole (24)

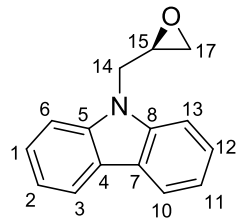


24

(±)-Epichlorohydrin was used to yield **24** as a racemic mixture (1.094 g, 41% yield).

M.p. 106 - 107 °C (lit. 113 °C)[176]; R_f (EtOAc/hexane 1:9) 0.26; δ_H (Acetone-d_6, 400 MHz) 2.59 (dd, *J* = 2.5, 5.1 Hz, 1H, H-17), 2.76 (dd, *J* = 4.0, 5.1 Hz, 1H, H-17), 3.34 (m, 1H, H-15), 4.42 (dd, *J* = 5.5, 15.9 Hz, 1H, H-14), 4.81 (dd, *J* = 3.0, 15.9 Hz, 1H, H-14), 7.22 (m, 2H, H-2, H-11), 7.45 (m, 2H, H-1, H-12), 7.63 (d, *J* = 8.3 Hz, 2H, H-6, H-13), 8.13 (m, 2H, H-3, H-10); δ_C (Acetone-d_6, 100.6 MHz) 45.4 (C_H), 45.6 (C_H), 51.3 (C_H), 110.4 (C_H), 120.1 (C_H), 121.0 (C_H), 123.8 (C_{quat}), 126.7 (C_H), 141.9 (C_{quat}).

(*S*)-9-(Oxiran-2-ylmethyl)-9*H*-carbazole (25)

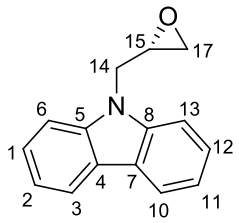


25

(*S*)-Epichlorohydrin was used to yield **25** as a pure enantiomer (1.346 g, 48% yield).

M.p. 73 - 74 °C; R_f (EtOAc/hexane 1:9) 0.26; δ_H (acetone-d_6, 400 MHz) 2.59 (dd, *J* = 2.5, 5.1 Hz, 1H, H-17), 2.76 (dd, *J* = 4.0, 5.1 Hz, 1H, H-17), 3.34 (m, 1H, H-15), 4.42 (dd, *J* = 5.5, 15.9 Hz, 1H, H-14), 4.81 (dd, *J* = 3.0, 15.9 Hz, 1H, H-14), 7.22 (m, 2H, H-2, H-11), 7.45 (m, 2H, H-1, H-12), 7.63 (d, *J* = 8.3 Hz, 2H, H-6, H-13), 8.13 (m, 2H, H-3, H-10); δ_C (acetone-d_6, 100.6 MHz) 45.4 (C_H), 45.6 (C_H), 51.3 (C_H), 110.4 (C_H), 120.1 (C_H), 121.0 (C_H), 123.8 (C_{quat}), 126.7 (C_H), 141.9 (C_{quat}).

(*R)*-9-(Oxiran-2-ylmethyl)-9*H*-carbazole (26)



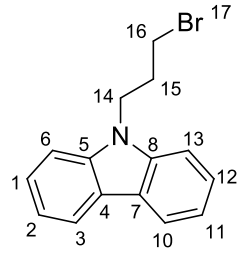
26

(*R*)-Epichlorohydrin was used to yield **26** as a pure enantiomer (1.292 g, 50% yield).

M.p. 81 - 82 °C; R_f (EtOAc/hexane 1:9) 0.26; δ_H (acetone-d_6, 400 MHz) 2.59 (dd, *J* = 2.5, 5.1 Hz, 1H, H-17), 2.76 (dd, *J* = 4.0, 5.1 Hz, 1H, H-17), 3.34 (m, 1H, H-15), 4.42 (dd, *J* = 5.5, 15.9 Hz, 1H, H-14), 4.81 (dd, *J* = 3.0, 15.9 Hz, 1H, H-14), 7.22 (m, 2H, H-2, H-11), 7.45 (m, 2H, H-1, H-12), 7.63 (d, *J* = 8.3 Hz, 2H, H-6, H-13), 8.13 (m, 2H, H-3, H-10); δ_C (Acetone-d_6, 100.6 MHz) 45.4 (C_H), 45.6 (C_H), 51.3 (C_H), 110.4 (C_H), 120.1 (C_H), 121.0 (C_H), 123.8 (C_{quat}), 126.7 (C_H), 141.9 (C_{quat}).

7.3.2 Miscellaneous Precursors

9-(3-Bromopropyl)-9*H*-carbazole (27)



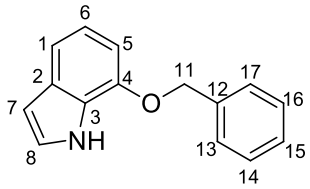
27

To a solution of NaH (271 mg, 6.78 mmols) and anhydrous THF (3 mL) at 0 °C, was added a solution of carbazole (1.50 g, 8.97 mmols) in THF (10 mL) dropwise over 20 min. The reaction mixture was warmed to room temperature and stirred for 1 hr to allow for complete deprotonation. 1,3-Dibromopropane (1.37 mL, 13.46 mmols, 1.5 eq.) was added to the reaction mixture and allowed to stir for 16 hr at room temperature. The reaction mixture was quenched with a saturated NH_4Cl solution and the combined organic components were extracted using EtOAc (2x 20 mL). The combined extracts were dried over anhydrous Na_2SO_4

and the crude product purified by column chromatography using 100% hexane to yield **27** as a white solid (404 mg, 16% yield).

M.p. 149 - 150 °C (lit. 148 – 150 °C)[177]; R_f (EtOAc/hexane 1:9) 0.48; δ_H (DMSO-d_6, 300 MHz) 2.31 (p, *J* = 6.9 Hz, 2H, H-15), 3.52 (t, *J* = 6.6 Hz, 2H, H-16), 4.51 (t, *J* = 6.9 Hz, 2H, H-14), 7.17-7.24 (m, 2 H, H-2, H-11), 7.42-7.50 (m, 2 H, H-1, H-12), 7.63 (d, *J* = 8.3 Hz, 2H, H-6, H-13), 8.16 (d, *J* = 7.7 Hz, 2H, H-3, H-10).

7-(Benzyloxy)-1*H*-indole (28)



28

To a 0 °C solution of NaH (126 mg, 3.15 mmols) and anhydrous THF (2 mL) was added a solution of 7-hydroxyindole (500 mg, 3.76 mmols) in THF (5 mL) drop wise over 20 min. The reaction mixture was warmed to room temperature and stirred for 1 hr to allow for complete deprotonation. Benzyl bromide (0.58 mL, 4.88 mmols) was added to the reaction mixture and stirred for 12 hr. The reaction mixture was quenched with a saturated NH_4Cl solution and the combined organic components were extracted using EtOAc (2x 20 mL). The combined extracts were dried over Na_2SO_4 and the crude product purified by column chromatography using EtOAc/ hexane (0:100 to 2:8) mixtures to yield **28** as a brown solid (373 mg, 44% yield).

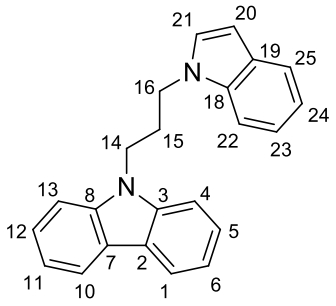
M.p. 72 - 73 °C (lit. 73 °C)[178]; R_f (EtOAc/Hexane 3:8) 0.67; δ_H (chloroform-*d*, 300 MHz) 5.19 (s, 2H, H-11), 6.51 (m, 1H, H-7), 6.70 (d, *J* = 7.7 Hz, 1H, H-5), 6.99 (t, *J* = 7.8 Hz, 1H, H-6), 7.12-7.15 (m, 1H, H-8), 7.20-7.48 (m, 6H, H-1, H-13, H-14, H-15, H-16, H-17), 8.38 (s, 1H, NH).

7.3.3 Procedure for Preparation of Carbazole Indoles 29-32

To a solution of NaH (42 mg, 1.05 mmols) in anhydrous THF (2 mL) was added drop wise a solution of Indole (122 mg, 1.04 mmols) in THF (4 mL) at 0 °C. The mixture was warmed to room temperature and allowed to stir for 1 hr. The appropriate epoxide or bromide derivative (1.04 mmols, 1 eq.) was added and the mixture was refluxed at 70 °C for 12 hr. The reaction mixture was cooled, quenched using a saturated NH_4Cl solution (2 mL) and then reduced

under reduced pressure to remove the residual THF. The combined organic components were dissolved in EtOAc and extracted (3 x 15 mL). The organic mixture was dried over Na_2SO_4 and purified by column chromatography using EtOAc / hexane (2:98 to 20:80) mixtures. The resulting solids were recrystallized from pentane.

9-(3-(1*H*-Indol-1-yl)propyl)-9*H*-carbazole (29)

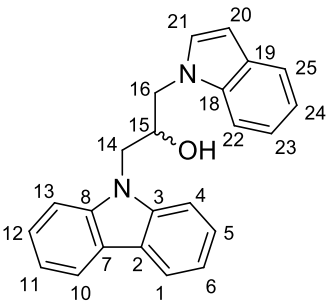


29

27 was used to yield **29** as a white solid (172 mg, 51% yield).

M.p. 65 - 66 °C; R_f (EtOAc/hexane 2:8) 0.55; δ_H (chloroform-*d*, 400 MHz) 2.45-2.53 (p, *J* = 7.1 Hz, 2H, H-15), 4.18-4.23 (t, *J* = 7.0 Hz, 2H, H-14/H16), 4.31-4.36 (t, *J* = 7.2 Hz, 2H, H-14/H-16), 6.58 (dd, *J* = 0.6, 3.2 Hz, 1H, H-20), 7.11 (d, *J* = 3.2 Hz, H-21), 7.12-7.18 (m, 1H, H_{Ar}), 7.19-7.28 (m, 6H, H_{Ar}), 7.44 (m, 2H, H_{Ar}), 7.69 (m, 1H, H_{Ar}) 8.13 (m, 2H, H_{Ar}); δ_C (chloroform-*d*, 100.6 MHz) 29.4 (C-15), 40.5 (C-14/C-16), 44.0 (C-14/C-16), 101.9 (C_H), 108.6 (C_H), 109.4 (C_H), 119.3 (C_H), 119.7 (C_H), 120.6 (C_H), 121.3 (C_H), 121.9 (C_H), 123.3 (C_{quat}), 126.0 (C_H), 127.5 (C_H), 129.0 (C_{quat}), 136.2 (C_{quat}), 140.4 (C_{quat}); HRMS-ES$^+$ Calculated: 325.1705 $[M+H]^+$ for $C_{23}H_{21}N_2$, Observed: 325.1693; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 99.7%.

1-(9*H*-Carbazol-9-yl)-3-(1*H*-indol-1-yl)propan-2-ol (30)

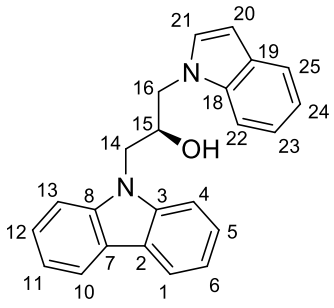


30

Racemic **24** was used to yield **30** as a white solid (192 mg, 54% yield).

ATR-FTIR ν_{max} /cm^{-1} 3538 (OH); M.p. 51 - 52 °C; R_f (EtOAc/hexane 2:8) 0.35; δ_H (chloroform-*d*, 400 MHz) 1.99 (s, 1H, OH) 4.25-4.32 (m, 2H, H-14/H-16), 4.33-4.46 (m, 2H, H-14/H-16), 4.56-4.65 (m, 1H, H-15), 6.57 (dd, *J* = 0.76, 3.2 Hz, 1H, H-20), 7.12-7.17 (m, 2H, H_{Ar}), 7.18-7.23 (m, 1H, H_{Ar}), 7.23-7.30 (m, 3H, H_{Ar}), 7.34 (m, 2H, H_{Ar}), 7.44-7.49 (m, 2H, H_{Ar}), 7.67 (m, 1H, H_{Ar}), 8.12 (m, 2H, H_{Ar}); δ_C (chloroform-*d*, 100.6 MHz) 47.2 (C-14/C-16), 50.4 (C-14/C-16), 70.6 (C_H), 102.4 (C_H), 109.0 (C_H), 109.4 (C_H), 119.7 (C_H), 120.0 (C_H), 120.6 (C_H), 121.4 (C_H), 122.1 (C_H), 123.4 (C_{quat}), 126.2 (C_H), 128.5 (C_H), 129.0 (C_{quat}), 136.4 (C_{quat}), 140.8 (C_{quat}); HRMS-ES$^+$ Calculated: 341.1654 $[M+H]^+$ for $C_{23}H_{21}N_2O$, Observed: 341.1645; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 98.0%; Chiral OD HPLC, flow rate 1.0 mL/min, *i*-PrOH / hexane (20:80), $^{\tau}$Major: 17.8 min, $^{\tau}$Minor: 23.1 min.

(*R*)-1-(9*H*-Carbazol-9-yl)-3-(1*H*-indol-1-yl)propan-2-ol (31)

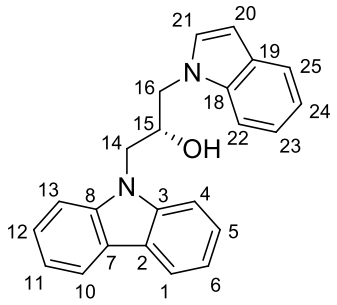


31

26 [(*R*)-isomer] was used to yield **31** as a white solid (252 mg, 72% yield).

ATR-FTIR ν_{max} /cm^{-1} 3538 (OH); M.p. 53 - 54 °C; R_f (EtOAc/hexane 2:8) 0.35; δ_H (chloroform-*d*, 400 MHz) 1.99 (s, 1H, OH) 4.25-4.32 (m, 2H, H-14/H-16), 4.33-4.46 (m, 2H, H-14/H-16), 4.56-4.65 (m, 1H, H-15), 6.57 (dd, *J* = 0.76, 3.2 Hz, 1H, H-20), 7.12-7.17 (m, 2H, H_{Ar}), 7.18-7.23 (m, 1H, H_{Ar}), 7.23-7.30 (m, 3H, H_{Ar}), 7.34 (m, 2H, H_{Ar}), 7.44-7.49 (m, 2H, H_{Ar}), 7.67 (m, 1H, H_{Ar}), 8.12 (m, 2H, H_{Ar}); δ_C (chloroform-*d*, 100.6 MHz) 47.2 (C-14/C-16), 50.4 (C-14/C-16), 70.6 (C_H), 102.4 (C_H), 109.0 (C_H), 109.4 (C_H), 119.7 (C_H), 120.0 (C_H), 120.6 (C_H), 121.4 (C_H), 122.1 (C_H), 123.4 (C_{quat}), 126.2 (C_H), 128.5 (C_H), 129.0 (C_{quat}), 136.4 (C_{quat}), 140.8 (C_{quat}); HRMS-ES$^+$ Calculated: 341.1654 $[M+H]^+$ for $C_{23}H_{21}N_2O$, Observed: 341.1637; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 95.9%; Chiral OD HPLC, flow rate 1.0 mL/min, *i*-PrOH / hexane (20:80), $^{\tau}$Major: 17.8 min, $^{\tau}$Minor: 23.7 min.

(*S*)-1-(9*H*-Carbazol-9-yl)-3-(1*H*-indol-1-yl)propan-2-ol (32)



32

25 [(*S*)-isomer] was used to yield **32** as a white solid (248 mg, 70% yield).

ATR-FTIR v_{max} /cm^{-1} 3538 (OH); M.p. 52 - 53 °C; R_f (EtOAc/hexane 2:8) 0.35; δ_H (chloroform-*d*, 400 MHz) 1.99 (s, 1H, OH) 4.25-4.32 (m, 2H, H-14/H-16), 4.33-4.46 (m, 2H, H-14/H-16), 4.56-4.65 (m, 1H, H-15), 6.57 (dd, *J* = 0.76, 3.2 Hz, 1H, H-20), 7.12-7.17 (m, 2H, H_{Ar}), 7.18-7.23 (m, 1H, H_{Ar}), 7.23-7.30 (m, 3H, H_{Ar}), 7.34 (m, 2H, H_{Ar}), 7.44-7.49 (m, 2H, H_{Ar}), 7.67 (m, 1H, H_{Ar}), 8.12 (m, 2H, H_{Ar}); δ_C (chloroform-*d*, 100.6 MHz) 47.2 (C-14/C-16), 50.4 (C-14/C-16), 70.6 (C_H), 102.4 (C_H), 109.0 (C_H), 109.4 (C_H), 119.7 (C_H), 120.0 (C_H), 120.6 (C_H), 121.4 (C_H), 122.1 (C_H), 123.4 (C_{quat}), 126.2 (C_H), 128.5 (C_H), 129.0 (C_{quat}), 136.4 (C_{quat}), 140.8 (C_{quat}); HRMS-ES^+ Calculated: 341.1654 $[M+H]^+$ for $C_{23}H_{21}N_2O$, Observed: 341.1639; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 93.0%; Chiral OD HPLC, flow rate 1.0 mL/min, *i*-PrOH / hexane (20:80), $^{\tau}$Major: 22.4 min, $^{\tau}$Minor: 17.8 min.

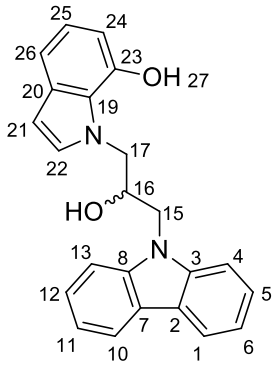
7.3.4 General Procedure for Carbazole Indoles 33-35

To a solution of NaH (40 mg, 1.0 mmol) in anhydrous NMP (1 mL) was added drop wise a solution of **28** (223 mg, 1.0 mmol) in NMP (2 mL) at 0 °C. The mixture was warmed to room temperature and allowed to stir for 1 hr. The appropriate epoxide (288 mg, 1.0 mmol) was added and the mixture was refluxed at 130 °C for 16 hr. The reaction mixture was cooled, quenched using a saturated NH_4Cl solution and extracted using EtOAc (3 x 15 mL). The organic mixture was washed with H_2O (9x 10 mL) and brine (10 mL) to remove the NMP, dried over Na_2SO_4 and purified by column chromatography with EtOAc / hexane (2:98 to 30:70) mixtures.

The resulting solid was dissolved in anhydrous MeOH (4 mL) in a reaction vessel flushed with nitrogen and 10 mol% of Pd-C (90 mg, 0.76 mmols) was added. A balloon filled with hydrogen was fitted to the vessel and the solution was stirred for 12 hr at room temperature. The

reaction was filtered through Celite© and reduced under reduced pressure to remove the excess solvent. The resulting solids were purified by column chromatography using EtOAc / hexane (2:8 to 3:7) mixtures and recrystallized using DCM.

1-(3-(9*H*-Carbazol-9-yl)-2-hydroxypropyl)-1*H*-indol-7-ol (33)

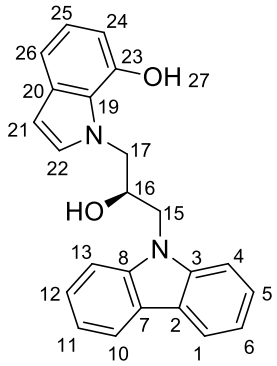


33

Racemic **24** was used to yield **33** as a white solid (123 mg, 35% yield).

ATR-FTIR ν_{max} /cm^{-1} 3457 (OH), 3218 (OH); M.p. 198 - 199 °C; R_f (EtOAc/hexane 2:8) 0.15; δ_H (methanol-d_4, 400 MHz) 4.29-4.35 (dd, *J* = 3.2, 15.0 Hz, 1H, H-15), 4.40-4.45 (m, 1H, H-15) 4.46-4.51 (m, 1H, H-17), 4.60-4.68 (m, 1H, H-16), 4.90 (m, 1H, H-17), 6.46 (d, *J* = 3.1 Hz, 1H, H-21), 6.60 (dd, *J* = 0.9, 7.8 Hz, 1H, H-24) 6.87-6.92 (t, *J* = 7.8 Hz, 1H, H-25) 7.13 (dd, *J* = 0.9, 7.8 Hz, 1H, H-26) 7.19-7.24 (m, 3H, H-6, H-11, H-22) 7.38-7.44 (m, 4H, H-4, H-5, H-12, H-13) 8.10 (m, 2H, H-1, H-10); δ_C (methanol-d_4, 100.6 MHz) 48.2 (C_H), 53.9 (C_H), 72.9 (C_H), 102.2 (C_H), 107.7 (C_H), 110.3 (C_H), 113.4 (C_H), 119.8 (C_H), 120.8 (C_H), 121.1 (C_H), 124.2 (C_{quat}), 126.6 (C_H), 127.0 (C_{quat}), 131.0 (C_H), 133.2 (C_{quat}), 142.3 (C_{quat}), 145.6 (C_{quat}); HRMS-ES$^+$ Calculated: 357.1603 $[M+H]^+$ for $C_{23}H_{21}N_2O_2$, Observed: 357.1588; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 96.3%; Chiral OD HPLC, flow rate 0.9 mL/min, *i*-PrOH / hexane (25:75), $^{\tau}$Major: 11.9 min, $^{\tau}$ Minor: 20.2 min.

(*R*)-1-(3-(9*H*-Carbazol-9-yl)-2-hydroxypropyl)-1*H*-indol-7-ol (34)

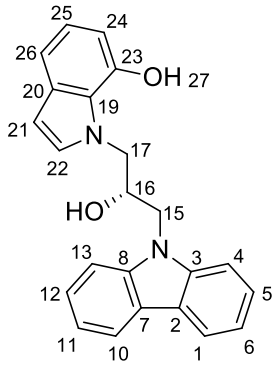


34

25 [(*S*)-isomer] was used to yield **34** as a white solid (141 mg, 40% yield).

ATR-FTIR v_{max} /cm^{-1} 3538 (OH), 3250 (OH); M.p. 175 - 176 °C; R_f (EtOAc/hexane 2:8) 0.15; δ_H (methanol-d_4, 400 MHz) 4.29-4.35 (dd, *J* = 3.2, 15.0 Hz, 1H, H-15), 4.40-4.45 (m, 1H, H-15) 4.46-4.51 (m, 1H, H-17), 4.60-4.68 (m, 1H, H-16), 4.90 (m, 1H, H-17), 6.46 (d, *J* = 3.1 Hz, 1H, H-21), 6.60 (dd, *J* = 0.9, 7.8 Hz, 1H, H-24) 6.87-6.92 (t, *J* = 7.8 Hz, 1H, H-25) 7.13 (dd, *J* = 0.9, 7.8 Hz, 1H, H-26) 7.19-7.24 (m, 3H, H-6, H-11, H-22) 7.38-7.44 (m, 4H, H-4, H-5, H-12, H-13) 8.10 (m, 2H, H-1, H-10); δ_C (methanol-d_4, 100.6 MHz) 48.2 (C_H), 53.9 (C_H), 72.9 (C_H), 102.2 (C_H), 107.7 (C_H), 110.3 (C_H), 113.4 (C_H), 119.8 (C_H), 120.8 (C_H), 121.1 (C_H), 124.2 (C_{quat}), 126.6 (C_H), 127.0 (C_{quat}), 131.0 (C_H), 133.2 (C_{quat}), 142.3 (C_{quat}), 145.6 (C_{quat}); HRMS-ES$^+$ Calculated: 357.1603 $[M+H]^+$ for $C_{23}H_{21}N_2O_2$, Observed: 357.1591; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 95.7%; Chiral OD HPLC, flow rate 0.9 mL/min, *i*-PrOH / hexane (25:75), $^{\tau}$Major: 11.9 min, $^{\tau}$Minor: 20.6 min.

(*S*)-1-(3-(9*H*-Carbazol-9-yl)-2-hydroxypropyl)-1*H*-indol-7-ol (35)



35

26 [(*R*)-isomer] was used to yield **35** as a white solid (162 mg, 45% yield).

ATR-FTIR v_{max} /cm^{-1} 3538 (OH), 3236 (OH); M.p. 193 - 194 °C; R_f (EtOAc/hexane 2:8) 0.15; δ_H (methanol-d_4, 400 MHz) 4.29-4.35 (dd, *J* = 3.2, 15.0 Hz, 1H, H-15), 4.40-4.45 (m, 1H, H-15) 4.46-4.51 (m, 1H, H-17), 4.60-4.68 (m, 1H, H-16), 4.90 (m, 1H, H-17), 6.46 (d, *J* = 3.1 Hz, 1H, H-21), 6.60 (dd, *J* = 0.9, 7.8 Hz, 1H, H-24) 6.87-6.92 (t, *J* = 7.8 Hz, 1H, H-25) 7.13 (dd, *J* = 0.9, 7.8 Hz, 1H, H-26) 7.19-7.24 (m, 3H, H-6, H-11, H-22) 7.38-7.44 (m, 4H, H-4, H-5, H-12, H-13) 8.10 (m, 2H, H-1, H-10); δ_C (methanol-d_4, 100.6 MHz) 48.2 (C_H), 53.9 (C_H), 72.9 (C_H), 102.2 (C_H), 107.7 (C_H), 110.3 (C_H), 113.4 (C_H), 119.8 (C_H), 120.8 (C_H), 121.1 (C_H), 124.2 (C_{quat}), 126.6 (C_H), 127.0 (C_{quat}), 131.0 (C_H), 133.2 (C_{quat}), 142.3 (C_{quat}), 145.6 (C_{quat}); HRMS-ES$^+$ Calculated: 357.1603 $[M+H]^+$ for $C_{23}H_{21}N_2O_2$, Observed: 357.1598; C18 HPLC, flow rate: 1 mL/min, H_2O / acetonitrile (30:70), 96.4%; Chiral OD HPLC, flow rate 0.9 mL/min, *i*-PrOH / hexane (25:75), $^{\tau}$ Major: 20.0 min, $^{\tau}$ Minor: 12.0 min.

7.4 General Analytical Methods

All reagents used were purchased from commercial sources (Fluka, Kimix and Sigma Aldrich). Double distilled deionised Millipore© Direct-Q water was used for all experiments. Eppendorf single and multi-channel micropipettes were used to deliver all micro-volumes manually. All pH measurements were carried out using a Crison 2000 MicropH meter which was calibrated using standard phosphate buffer solutions (pH 7.00 ± 0.02 and 4.00 ± 0.02) before use.

7.5 NP-40 Mediated β-Haematin Inhibition Assay

The β-haematin inhibition assay as described by Carter *et al* [92] was modified for manual liquid delivery. Test compounds and standards were dissolved in DMSO to make up 20 mM stock solutions and 20 µL of each solution was added to column 12 of a 96-well plate together with AR grade acetone (70 µL), NP40 (40 µL , 305.5 µM) and water (70 µL). A solution containing water/ acetone/ NP40 (305.5 µM) / DMSO at a v/v ratio of 35%/35%/20%/10% was made up and 100 µL was added to columns 1 to 11. A serial dilution from column 12 to column 2 was carried out leaving column 1 as the blank column. A solution of haematin (25mM) was prepared by sonicating haemin in DMSO for 1 min and then suspending 178.8 µL in 20 mL of a previously prepared 1M acetate buffer (pH 4.8). 100 µL of this haematin solution was added to all 96 wells. The plate was covered and incubated in an incubator at 37 °C for 5-6 hr. The plate was analysed using the pyridine-ferrichrome method developed by Ncokazi and Egan.[67] A solution of pyridine/ water/ acetone/ 2M HEPES buffer (pH 7.4) in ratios (v/v) of 50%/ 20%/ 20%/ 10% was prepared and 32 µL was added to each well. Lastly 60 µL of acetone was added to all wells to assist in the dispersion of haematin. The 96-well plate was shaken by mechanical means followed by the UV-visible absorbance being read using a SpectraMax P340 plate reader. Sigmoidal dose-response curves were plotted using the absorbance data together with Graphpad Prism v4.00. [179]

7.6 Turbidimetric Solubility Assay

A phosphate-buffered saline (PBS) (0.01M, pH 7.4) solution was prepared using five intact PBS tablets in 1 L of H_2O. Test compounds and hydrocortisone (standard) were dissolved in DMSO to make up 10 mM solutions and used to prepare a pre-dilution plate by serial dilution of the compounds in triplicate, in a 96-well plate, to yield concentrations from 8.0 mM to 0.25 mM. A final 96-well plate was prepared by pipetting 196 µL of DMSO in half of the wells in the plate and 196 µL of the previously prepared 0.01M PBS buffer. 4 µL of each solution in the pre-dilution plate was pipetted in triplicate into both the DMSO and PBS buffer containing wells for a final volume of 200 µL and concentration range of 200 µM to 0 µM (Figure 7.1). The plate was covered and left to equilibrate for 2 hr at ambient temperature.

	DMSO						0.01M pH 7.4 PBS					
	Compound 1 (triplicate)			Compound 2 (triplicate)			Compound 1 (triplicate)			Compound 2 (triplicate)		
Conc. (µM)	1	2	3	1	2	3	1	2	3	1	2	3
0.0												
5.0												
10.0												
20.0												
40.0												
80.0												
160.0												
200.0												

Figure 7.1: 96-Well plate layout

The UV-visible absorbance of the 96-well plate was read on a SpectraMax P340 plate reader and the corrected absorbance was calculated using the following formula: Absorbance at conc X – Blank absorbance = corrected absorbance for all wells. Plots of absorbance against compound concentration were prepared using MS excel 2013.

7.7 Computational Methods

7.7.1 Bayesian Modelling

All structures were drawn in Marvin sketch [180] and imported into Discovery Studio v4.0 [169] where all Bayesian statistic calculations was carried out using and the Bayesian models described by Wicht *et al*.[127] All structural data was analysed using the predetermined cut-offs of the relevant Bayesian models.

7.7.2 Docking Studies

All docking studies was carried out on Materials Studio 2016 version 16.1.0.21[162] using preformed β-haematin crystal structure faces ((001) and (011)) and an optimised force-field specifically designed for this system developed by the de Villiers group at Stellenbosch University (unpublished work). Compounds were drawn in Chemdraw Ultra 12.0.2[181],

imported to Material Studio and optimised using a Forcite geometry optimisation calculation (quality: Ultra-fine, Charge: charge using Qeq and max iterations: 20000). These optimised compounds were adsorbed onto both the β-haematin crystal structure faces using the adsorption locator (5 cycles, 70000 steps per cycle, current charge). The lowest energy adsorption complexes that were plausible were subjected to a Forcite anneal calculation (quality: ultra-fine, annealing cycles: 100, Charge: use current). The lowest plausible energy configuration of the complex was lastly subjected to a Forcite energy calculation (quality: ultra-fine, Charge: use current). The adsorption energy of the Inhibitor was calculated according to equation 6.1:

$$E_{ads}\ (kcal.mol^{-1}) = E_{final} - E_{initial}$$

$$= E\ (crystal + inhibitor) - E\ (crystal) - E\ (inhibitor) \qquad 6.1$$

References

(1) Hempelmann, E.; Krafts, K. **2013**, 0–13.

(2) Cox, F. E. *Parasites and Vectors* **2010**, *3* (1), 1-9.

(3) Kean, B. H.; Mott, K. E.; Adair, J. *Rev. Infect. Dis.* **1982**, *4* (4), 908–911.

(4) Ross, R.; Smyth, J. *Br. Med. J.* **1897**, *2* (1929), 1786–1788.

(5) Perkins, S. L.; Austin, C. C. *J. Parasitol* **2017**, *95* (2), 424–433.

(6) Foley, M. *Pharmacol. Ther.* **1998**, *79* (1), 55–87.

(7) *World Health Organisation Malaria Report 2017*; Switzerland, 2017.

(8) Greenwood, B. M.; Fidock, D. A.; Kyle, D. E.; IKappe, S. H. *Www . Jci . Org* **2008**, *118* (4), 1266-1276.

(9) Newton, BM, MRCP, P.; White, DSc, FRCP, N. *Annu. Rev. Med.* **1999**, *50* (1), 179–192.

(10) Bruce-chwatt, L. J. *Bull. World Health Organ.* **1965**, *32* (1887), 363–387.

(11) Goldberg, D. E.; Slater, A. F.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (8), 2931–2935.

(12) Egan, T. J. *Mol. Biochem. Parasitol.* **2008**, *157* (2), 127–136.

(13) Lew, V. L.; Tiffert, T.; Ginsburg, H. *Blood* **2003**, *101* (10), 4189-4194.

(14) Krugliak, M.; Zhang, J.; Ginsburg, H. *Mol. Biochem. Parasitol.* **2002**, *119* (2), 249–256.

(15) Egan, T. J. *J. Inorg. Biochem.* **2008**, *102* (5–6), 1288–1299.

(16) Na-Bangchang, K.; Karbwang, J. *Fundam. Clin. Pharmacol.* **2009**, *23* (4), 387–409.

(17) Bruce-Chwatt LJ. *World Health Organization, Monograph Series. No. 27*. 1986, p 261.

(18) Plowe, C. V; Cortese, J. F.; Djimde, A.; Nwanyanwu, O. C.; Watkins, W. M.; Winstanley, P. A.; Estrada-franco, J. G.; Mollinedo, R. E.; Avila, J. C.; Cespedes, J. L.; Carter, D.; Doumbo, O. K. *J. Infect. Dis.* **2018**, No. January, 1590–1596.

(19) Nirmalan, N.; Wang, P.; Sims, P. F. G.; Hyde, J. E. *Mol. Microbiol.* **2002**, *46* (1), 179–190.

(20) Chopra, I.; Roberts, M. *Microbiol. Mol. Biol. Rev.* **2001**, *65* (2), 232–260.

(21) Srivastava, I. K.; Rottenberg, H.; Vaidya, A. B. *J. Biol. Chem.* **1997**, *272* (7), 3961–3966.

(22) Looareesuwan, S.; Viravan, C.; Webster, H. K.; Kyle, D. E.; Hutchinson, D. B.; Canfield, C. J. *Am. J. Trop. Med. Hyg.* **1996**, *54* (1), 62–66.

(23) Davey, D. G. *British Medical Bulletin*. 1951, pp 37–46.

(24) Taylor, W. R. J.; Rigal, J.; Olliaro, P. L. *J. Vector Borne Dis.* **2003**, *40* (3–4), 65–72.

(25) Abdulla, S.; Ashley, E. A.; Bassat, Q.; Bethell, D.; Björkman, A.; Borrmann, S.; D'Alessandro, U.; Dahal, P.; Day, N. P.; Diakite, M.; Djimde, A. A.; Dondorp, A. M.;

Duong, S.; Edstein, M. D.; Fairhurst, R. M.; Faiz, M. A.; Falade, C.; Flegg, J. A.; Fogg, C.; Gonzalez, R.; Greenwood, B.; Guérin, P. J.; Guthmann, J. P.; Hamed, K.; Hien, T. T.; Htut, Y.; Juma, E.; Lim, P.; Mårtensson, A.; Mayxay, M.; Mokuolu, O. A.; Moreira, C.; Newton, P.; Noedl, H.; Nosten, F.; Ogutu, B. R.; Onyamboko, M. A.; Owusu-Agyei, S.; Phyo, A. P.; Premji, Z.; Price, R. N.; Pukrittayakamee, S.; Ramharter, M.; Sagara, I.; Se, Y.; Suon, S.; Stepniewska, K.; Ward, S. A.; White, N. J.; Winstanley, P. A. *Malar. J.* **2015**, *14*, 359, 1-12.

(26) Harmse, R.; Wong, H. N.; Smit, F.; Haynes, R. K.; N'Da, D. D. *Curr. Med. Chem.* **2015**, *22* (31), 3607–3630.

(27) Burrows, P. D. *Nat. Med.* **2012**, *13* (7), 625–627.

(28) O'Neill, P. M.; Barton, V. E.; Ward, S. A. *Molecules* **2010**, *15* (3), 1705–1721.

(29) Kamchonwongpaisan, S.; Meshnick, S. R. *Gen. Pharmacol.* **1996**, *27* (4), 587–592.

(30) Haber, F.; Weiss, J. *R. Society* **1934**, *147* (861), 332–351.

(31) Faurant, C. *Parasite* **2011**, *18* (3), 215–218.

(32) O'Neill, P. M.; Bray, P. G.; Hawley, S. R.; Ward, S. A.; Park, B. K. *Pharmacol. Ther.* **1998**, *77* (1), 29–58.

(33) Ridley, R. G. *Nature* **2002**, *415* (6872), 686–693.

(34) C. E. Song. In *Cinchona Alkaloids in Sythesis and Catalysis: Ligands, Immobilization and Organocatalysis*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2009.

(35) Berliner, R. W.; Earle, D. P.; Taggart, J. V; Welch, W. J.; Zubrod, C. G.; Knowlton, P.; Atchley, J. a; Shannon, J. a. *J. Clin. Invest.* **1948**, *27* (3 Pt 2), 108–113.

(36) Pou, S.; Winter, R. W.; Nilsen, A.; Kelly, J. X.; Li, Y.; Doggett, J. S.; Riscoe, E. W.; Wegmann, K. W.; Hinrichs, D. J.; Riscoe, M. K. *Antimicrob. Agents Chemother.* **2012**, *56* (7), 3475–3480.

(37) Ohnmacht, C. J.; Patel, A. R.; Lutz, R. E. *J. Med. Chem.* **1971**, *14* (10), 926–928.

(38) Biswas, S.; Valecha, N.; Tyagi, P. K.; Phookan, S.; Dev, V. *J. Vect. Borne. Dis* **2003**, No. December, 92–99.

(39) Kremsner, P. G.; Winkler, S.; Brandts, C. **2018**, *169* (2), 467–470.

(40) Barnes, K. I.; White, N. J. *Acta Trop.* **2005**, *94* (3 SPEC. ISS.), 230–240.

(41) White, N. J. *J.Clin.Invest* **2004**, *113* (0021–9738 (Print)), 1084–1092.

(42) Sanchez, C. P.; Dave, A.; Stein, W. D.; Lanzer, M. *Int. J. Parasitol.* **2010**, *40* (10), 1109–1118.

(43) Cowman, A. F. *Int. J. Parasitol.* **2001**, *31* (9), 871–878.

(44) Wang, N.; Wicht, K. J.; Imai, K.; Wang, M. Q.; Anh Ngoc, T.; Kiguchi, R.; Kaiser, M.; Egan, T. J.; Inokuchi, T. *Bioorganic Med. Chem.* **2014**, *22* (9), 2629–2642.

(45) Noedl, H.; Se, Y.; Schaecher, K.; Smith, B. L.; Socheat, D.; Fukuda, M. M. *N. Engl. J.*

Med. **2008**, *359* (24), 2619–2620.

(46) Fitch, C. D.; Gonzalez, Y.; Chevli, R. *J. Pharmacol. Exp. Ther.* **1975**, *195* (3), 397–403.

(47) Krogstad, D. J.; Gluzman, I. Y.; Kyle, D. E.; Oduola, A. M. J.; Martin, S. K.; Milhous, W. K.; Schlesinger, P. H.; Schlesinger, P. H. *Science (80-.).* **1987**, *238* (4831), 1283–1285.

(48) Gligorijevic, B.; Bennett, T.; McAllister, R.; Urbach, J. S.; Roepe, P. D. *Biochemistry* **2006**, *45* (41), 12411–12423.

(49) Ginsburg, H.; Krugliak, M. *Biochem. Pharmacol.* **1992**, *43* (1), 63–70.

(50) Martiney, J. A.; Cerami, A.; Slater, A. F. G. *J. Biol. Chem.* **1995**, *270* (38), 22393–22398.

(51) Bray, P. G.; Mungthin, M.; Ridley, R. G.; Ward, S. A. *Mol. Pharmacol.* **1998**, *54* (1), 170–179.

(52) Fidock, D. A.; Nomura, T.; Talley, A. K.; Cooper, R. A.; Dzekunov, S. M.; Ferdig, M. T.; Ursos, L. M. B.; bir Singh Sidhu, A.; Naudé, B.; Deitsch, K. W.; Su, X.; Wootton, J. C.; Roepe, P. D.; Wellems, T. E. *Mol. Cell* **2000**, *6* (4), 861–871.

(53) Picot, S.; Olliaro, P.; de Monbrison, F.; Bienvenu, A.-L.; Price, R. N.; Ringwald, P. *Malar. J.* **2009**, *8* (1), 89, 1-15.

(54) Lekostaj, J. K.; Natarajan, J. K.; Paguio, M. F.; Wolf, C.; Roepe, P. D. *Biochemistry* **2008**, *47* (39), 10394–10406.

(55) Juge, N.; Moriyama, S.; Miyaji, T.; Kawakami, M.; Iwai, H.; Fukui, T.; Nelson, N.; Omote, H.; Moriyama, Y. *Proc. Natl. Acad. Sci.* **2015**, *112* (11), 3356–3361.

(56) Parker, F. S.; Irvin, J. L. *J. Biol. Chem.* **1952**, *199* (889).

(57) Cohen, S. N.; Yielding, K. L. **1965**, *54* (1950), 521–527.

(58) O'Brien, R. L.; Olenick, J. G. G.; Hahn, F. E.; Hahn, R. E. *Biochemistry* **1966**, *55* (6), 1511–1517.

(59) Fitch, C. D.; Kanjananggulpan, P. *J. Biol. Chem.* **1987**, *262* (32), 15552–15555.

(60) Slater, A. F.; Cerami, A. *Nature* **1992**, *355* (6356), 167–169.

(61) Dorn, A.; Vippagunta, S. R.; Matile, H.; Jaquet, C.; Vennerstrom, J. L.; Ridley, R. G. *Biochem. Pharmacol.* **1998**, *55* (6), 727–736.

(62) Dorn, A.; Vippagunta, S. R.; Matile, H.; Bubendorf, A.; Vennerstrom, J. L.; Ridley, R. G. *Biochem. Pharmacol.* **1998**, *55* (6), 737–747.

(63) Dorn, A.; Stoffel, R.; Matile, H.; Bubendorf, A.; Ridley, R. G. *Nature* **1995**, *374* (6519), 269–271.

(64) Vippagunta, S. R.; Dorn, A.; Matile, H.; Bhattacharjee, A. K.; Karle, J. M.; Ellis, W. Y.; Ridley, R. G.; Vennerstrom, J. L. *J. Med. Chem.* **1999**, *42* (22), 4630–4639.

(65) Kaschula, C. H.; Egan, T. J.; Hunter, R.; Basilico, N.; Parapini, S.; Taramelli, D.; Pasini, E.; Monti, D. *J. Med. Chem.* **2002**, *45* (16), 3531–3539.

(66) Parapini, S.; Basilico, N.; Pasini, E.; Egan, T. J.; Olliaro, P.; Taramelli, D.; Monti, D. *Exp.*

Parasitol. **2000**, *96* (4), 249–256.

(67) Ncokazi, K. K.; Egan, T. J. *Anal. Biochem.* **2005**, *338* (2), 306–319.

(68) Lu, W. J.; Wicht, K. J.; Wang, L.; Imai, K.; Mei, Z. W.; Kaiser, M.; El Sayed, I. E. T.; Egan, T. J.; Inokuchi, T. *Eur. J. Med. Chem.* **2013**, *64*, 498–511.

(69) Kelly, J. X.; Winter, R.; Peyton, D. H.; Hinrichs, D. J.; Riscoe, M. *Antimicrob. Agents Chemother.* **2002**, *46* (1), 144–150.

(70) Gligorijevic, B.; McAllister, R.; Urbach, J. S.; Roepe, P. D. *Biochemistry* **2006**, *45* (41), 12400–12410.

(71) Combrinck, J. M.; Mabotha, T. E.; Ncokazi, K. K.; Ambele, M. A.; Taylor, D.; Smith, P. J.; Hoppe, H. C.; Egan, T. J. *ACS Chem. Biol.* **2013**, *8* (1), 133–137.

(72) Brown, W. H. *J. Exp. Med.* **1911**, *13* (2), 290–299.

(73) Slater, A. F.; Swiggard, W. J.; Orton, B. R.; Flitter, W. D.; Goldberg, D. E.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci.* **1991**, *88* (2), 325–329.

(74) Bohle, D. S.; Debrunner, P.; Jordan, P. A.; Madsen, S. K.; Schulz, C. E. *J. Am. Chem. Soc.* **1998**, *120* (32), 8255–8256.

(75) Bohle, D. S.; Dinnebier, R. E.; Madsen, S. K.; Stephens, P. W. *J. Biol. Chem.* **1997**, *272* (2), 713–716.

(76) Pagola, S.; Stephens, P. W.; Bohle, D. S.; Kosar, a D.; Madsen, S. K. *Nature* **2000**, *404*, 307–310.

(77) Sullivan, D. J.; Gluzman, I. Y.; Russell, D. G.; Goldberg, D. E. *Proc. Natl. Acad. Sci.* **1996**, *93* (21), 11865–11870.

(78) Sullivan, D. J.; Gluzman, I. Y.; Goldberg, D. E. *Science* **1996**, *271* (5246), 219–222.

(79) Egan, T. J.; Ross, D. C.; Adams, P. A. *FEBS Lett.* **1994**, *352* (1), 54–57.

(80) Olafson, K. N.; Ketchum, M. A.; Rimer, J. D.; Vekilov, P. G. *Proc. Natl. Acad. Sci.* **2015**, *112* (16), 4946–4951.

(81) Pisciotta, J. M.; Coppens, I.; Tripathi, A. K.; Scholl, P. F.; Shuman, J.; Bajad, S.; Shulaev, V.; Sullivan, D. J. *Biochem. J.* **2007**, *402* (1), 197–204.

(82) Kuter, D.; Mohunlal, R.; Fitzroy, S.-M.; Asher, C.; Smith, P. J.; Egan, T. J.; de Villiers, K. A. *CrystEngComm* **2016**, *18* (27), 5177–5187.

(83) Kuter, D.; Benjamin, S. J.; Egan, T. J. *J. Inorg. Biochem.* **2014**, *133* (January), 40–49.

(84) Buller, R.; Peterson, M. L.; Almarsson, Ö.; Leiserowitz, L. *Cryst. Growth Des.* **2002**, *2* (6), 553–562.

(85) de Villiers, K. A.; Marques, H. M.; Egan, T. J. *J. Inorg. Biochem.* **2008**, *102* (8), 1660–1667.

(86) De Villiers, K. A.; Gildenhuys, J.; Le Roex, T. *ACS Chem. Biol.* **2012**, *7* (4), 666–671.

(87) Kurosawa, Y.; Dorn, A.; Kitsuji-shirane, M.; Satoh, T.; Matile, H.; Hofheinz, W.;

Masciadri, R.; Kansy, M.; Ridley, R. G.; Kurosawa, Y. A. E.; Shimada, H. *Antimicrob. Agents Chemother.* **2000**, *44* (10), 2638–2644.

(88) Fitch, C. D.; Cai, G. Z.; Chen, Y. F.; Shoemaker, J. D. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1999**, *1454* (1), 31–37.

(89) Dominguez, A.; Fernandez, A.; Gonzalez, N.; Iglesias, E.; Montenegro, L. *J. Chem. Educ.* **1997**, *74* (10), 1227-1231.

(90) Huy, N. T.; Uyen, D. T.; Maeda, A.; Trang, D. T. X.; Oida, T.; Harada, S.; Kamei, K. *Antimicrob. Agents Chemother.* **2007**, *51* (1), 350–353.

(91) Carter, M.; Phelan, V.; Sandlin, R.; Bachmann, B.; Wright, D. *Comb. Chem. High Throughput Screen.* **2010**, *13* (3), 285–292.

(92) Sandlin, R. D.; Carter, M. D.; Lee, P. J.; Auschwitz, J. M.; Leed, S. E.; Johnson, J. D.; Wright, D. W. *Antimicrob. Agents Chemother.* **2011**, *55* (7), 3363–3369.

(93) Plouffe, D.; Brinker, A.; McNamara, C.; Henson, K.; Kato, N.; Kuhen, K.; Nagle, A.; Adrián, F.; Matzen, J. T.; Anderson, P.; Nam, T.-G.; Gray, N. S.; Chatterjee, A.; Janes, J.; Yan, S. F.; Trager, R.; Caldwell, J. S.; Schultz, P. G.; Zhou, Y.; Winzeler, E. a. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (26), 9059–9064.

(94) Gamo, F.-J.; Sanz, L. M.; Vidal, J.; de Cozar, C.; Alvarez, E.; Lavandera, J.-L.; Vanderwall, D. E.; Green, D. V. S.; Kumar, V.; Hasan, S.; Brown, J. R.; Peishoff, C. E.; Cardon, L. R.; Garcia-Bustos, J. F. *Nature* **2010**, *465* (7296), 305–310.

(95) Guiguemde, W. A.; Shelat, A. A.; Garcia-Bustos, J. F.; Diagana, T. T.; Gamo, F. J.; Guy, R. K. *Chem. Biol.* **2012**, *19* (1), 116–129.

(96) Kotz, J. *Sci. Exch.* **2012**, *5* (15), 1–3.

(97) Rush, M. A.; Baniecki, M. L.; Mazitschek, R.; Cortese, J. F.; Wiegand, R.; Clardy, J.; Wirth, D. F. *Antimicrob. Agents Chemother.* **2009**, *53* (6), 2564–2568.

(98) Guiguemde, W. A.; Shelat, A. A.; Bouck, D.; Duffy, S.; Gregory, J.; Davis, P. H.; Smithson, D. C.; Connelly, M.; Clark, J.; Jiménez-díaz, M. B.; Martinez, M. S.; Wilson, E. B.; Tripathi, A. K.; Riscoe, M. K.; Phillips, M. A.; Rathod, P. K.; Van, W. C. *Nature* **2010**, *465* (7296), 311–315.

(99) Smilkstein, M.; Sriwilaijaroen, N.; Kelly, J. X.; Wilairat, P.; Riscoe, M. *Antimicrob. Agents Chemother.* **2004**, *48* (5), 1803–1806.

(100) Sandlin, R. D.; Fong, K. Y.; Wicht, K. J.; Carrell, H. M.; Egan, T. J.; Wright, D. W. *Int. J. Parasitol. Drugs Drug Resist.* **2014**, *4* (3), 316–325.

(101) Ladenberg A. (1875) Derivatives of diamines. Deut Chem Ges Ber;8:677–678. - Google Scholar https://scholar-google-co-za.ezproxy.uct.ac.za/scholar?q=+Ladenberg+A.+%281875%29+Derivatives+of+diamines.+Deut+Chem+Ges+Ber%3B8%3A677–678.&btnG=&hl=en&as_sdt=0%2C5 (accessed Aug 15, 2017).

(102) Hobrecker F. (1872) Reduction-products of nitracetamide compounds. Deut Chem Ges Ber;5:920–924 - Google Scholar https://scholar-google-co-

za.ezproxy.uct.ac.za/scholar?hl=en&q=+Hobrecker+F.+%281872%29+Reduction-products+of+nitracetamide+compounds.+Deut+Chem+Ges+Ber%3B5%3A920–924&btnG=&as_sdt=1%2C5&as_sdtp= (accessed Aug 15, 2017).

(103) Zhang, L.; Peng, X.-M.; Damu, G. L. V.; Geng, R.-X.; Zhou, C.-H. *Med. Res. Rev.* **2014**, *34* (2), 340–437.

(104) Ayhan-Kilcigil, G.; Kuş, C.; Özdamar, E. D.; Can-Eke, B.; Içan, M. *Arch. Pharm. (Weinheim).* **2007**, *340* (11), 607–611.

(105) Pabba, C.; Wang, H. J.; Mulligan, S. R.; Chen, Z. J.; Stark, T. M.; Gregg, B. T. *Tetrahedron Lett.* **2005**, *46* (44), 7553–7557.

(106) Fonseca, T.; Gigante, B.; Gilchrist, T. L. *Tetrahedron* **2001**, *57* (9), 1793–1799.

(107) Denny, W. A.; Rewcastle, G. W.; Baguley, B. C. *J. Med. Chem.* **1990**, *33*, 819–193.

(108) Mayence, A.; Vanden Eynde, J. J.; Kaiser, M.; Brun, R.; Yarlett, N.; Huang, T. L. *Bioorganic Med. Chem.* **2011**, *19* (24), 7493–7500.

(109) Miller, J. F.; Turner, E. M.; Gudmundsson, K. S.; Jenkinson, S.; Spaltenstein, A.; Thomson, M.; Wheelan, P. *Bioorganic Med. Chem. Lett.* **2010**, *20* (7), 2125–2128.

(110) Navarrete-Vázquez, G.; Hidalgo-Figueroa, S.; Torres-Piedra, M.; Vergara-Galicia, J.; Rivera-Leyva, J. C.; Estrada-Soto, S.; León-Rivera, I.; Aguilar-Guardarrama, B.; Rios-Gómez, Y.; Villalobos-Molina, R.; Ibarra-Barajas, M. *Bioorganic Med. Chem.* **2010**, *18* (11), 3985–3991.

(111) Achar, K. C. S.; Hosamani, K. M.; Seetharamareddy, H. R. *Eur. J. Med. Chem.* **2010**, *45* (5), 2048–2054.

(112) Dieckmann-Schuppert, A.; Franklin, R. M. *Antimicrob Agents Chemother* **1990**, *34* (8), 1529–1534.

(113) Skinner-Adams, T. S.; Davis, T. M. E.; Manning, L. S.; Johnston, W. A. *Trans. R. Soc. Trop. Med. Hyg.* **1997**, *91* (5), 580–584.

(114) Worachartcheewan, A.; Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Prachayasittikul, V. *EXCLI J.* **2013**, *12* (11), 569–583.

(115) Ramachandran, S.; Hameed P., S.; Srivastava, A.; Shanbhag, G.; Morayya, S.; Rautela, N.; Awasthy, D.; Kavanagh, S.; Bharath, S.; Reddy, J.; Panduga, V.; Prabhakar, K. R.; Saralaya, R.; Nanduri, R.; Raichurkar, A.; Menasinakai, S.; Achar, V.; Jiménez-Díaz, M. B.; Martínez, M. S.; Angulo-Barturen, I.; Ferrer, S.; Sanz, L. M.; Gamo, F. J.; Duffy, S.; Avery, V. M.; Waterson, D.; Lee, M. C. S.; Coburn-Flynn, O.; Fidock, D. A.; Iyer, P. S.; Narayanan, S.; Hosagrahara, V.; Sambandamurthy, V. K. *J. Med. Chem.* **2014**, *57* (15), 6642–6652.

(116) Singh, K.; Okombo, J.; Brunschwig, C.; Ndubi, F.; Barnard, L.; Wilkinson, C.; Njogu, P. M.; Njoroge, M.; Laing, L.; Machado, M.; Prudêncio, M.; Reader, J.; Botha, M.; Nondaba, S.; Birkholtz, L. M.; Lauterbach, S.; Churchyard, A.; Coetzer, T. L.; Burrows, J. N.; Yeates, C.; Denti, P.; Wiesner, L.; Egan, T. J.; Wittlin, S.; Chibale, K. *J. Med. Chem.* **2017**, *60* (4), 1432–1448.

(117) Gaba, M.; Mohan, C. *Development of drugs based on imidazole and benzimidazole bioactive heterocycles: Recent advances and future directions*; Springer US, 2016; Vol. 25.

(118) Knölker, H. J.; Reddy, K. R. *Chem. Rev.* **2002**, *102* (11), 4303–4427.

(119) Das, K. C.; Chakraborty, D. P.; Bose, P. K. *Experientia* **1965**, *21* (6), 340-340.

(120) Akué-Gédu, R.; Rossignol, E.; Azzaro, S.; Knapp, S.; Filippakopoulos, P.; Bullock, A. N.; Bain, J.; Cohen, P.; Prudhomme, M.; Anizon, F.; Moreau, P. *J. Med. Chem.* **2009**, *52* (20), 6369–6381.

(121) Zhu, G.; Conner, S. E.; Zhou, X.; Chan, H. K.; Shih, C.; Engler, T. A.; Al-Awar, R. S.; Brooks, H. B.; Watkins, S. A.; Spencer, C. D.; Schultz, R. M.; Dempsey, J. A.; Considine, E. L.; Patel, B. R.; Ogg, C. A.; Vasudevan, V.; Lytle, M. L. *Bioorganic Med. Chem. Lett.* **2004**, *14* (12), 3057–3061.

(122) Molette, J.; Routier, J.; Abla, N.; Besson, D.; Bombrun, A.; Brun, R.; Burt, H.; Georgi, K.; Kaiser, M.; Nwaka, S.; Muzerelle, M.; Scheer, A. *ACS Med. Chem. Lett.* **2013**, *4* (11), 1037–1041.

(123) Gu, W.; Qiao, C.; Wang, S.-F.; Hao, Y.; Miao, T.-T. *Bioorg. Med. Chem. Lett.* **2014**, *24* (1), 328–331.

(124) Bashir, M.; Bano, A.; Ijaz, A. S.; Chaudhary, B. A. *Molecules* **2015**, *20* (8), 13496–13517.

(125) Brown, V. H.; 2-Rad, M.; Dectraw, J. I. *J. Med. Chem.* **1971**, *14* (6), 549–550.

(126) Wang, T.; Mäser, P.; Picard, D. *J. Med. Chem.* **2016**, *59* (13), 6344–6352.

(127) Wicht, K. J.; Combrinck, J. M.; Smith, P. J.; Egan, T. J. *Bioorganic Med. Chem.* **2015**, *23* (16), 5210–5217.

(128) Crisman, T. J.; Bender, A.; Milik, M.; Jenkins, J. L.; Scheiber, J.; Sukuru, S. C. K.; Fejzo, J.; Hommel, U.; Davies, J. W.; Glick, M. *J. Med. Chem.* **2008**, *51* (8), 2481–2491.

(129) Klon, A. E.; Glick, M.; Thoma, M.; Acklin, P.; Davies, J. W. *J. Med. Chem.* **2004**, *47* (11), 2743–2749.

(130) Diller, D. J.; Hobbs, D. W. *J. Med. Chem.* **2004**, *47* (25), 6373–6383.

(131) Fang, J.; Yang, R.; Gao, L.; Zhou, D.; Yang, S.; Liu, A. L.; Du, G. H. *J. Chem. Inf. Model.* **2013**, *53* (11), 3009–3020.

(132) Balfer, J.; Bajorath, J. *J. Chem. Inf. Model.* **2014**, *54* (9), 2451–2468.

(133) Van Hoorn, W. P.; Bell, A. S. *J. Chem. Inf. Model.* **2009**, *49* (10), 2211–2220.

(134) Bayes, T. *Phil. Trans* **1963**, *53*, 370-418.

(135) Discovery Studio Bayesian Modelling, *Accelrys Software Inc.,* San Diego, http://accelrys.com/resource-center/case-studies/bayesian-modeling.html

(136) Bender, A.; Mussa, H. Y.; Glen, R. C. *J. Biomol. Screen.* **2005**, *10* (7), 658–666.

(137) Xia, X.; Maliski, E. G.; Gallant, P.; Rogers, D. *J.Med.Chem.* **2004**, *47*, 4463–4470.

(138) Liao, C.; Sitzmann, M.; Pugliese, A.; Nicklaus, M. C. *Future Med. Chem.* **2011**, *3* (8), 1057–1085.

(139) Anand, K.; Wakode, S. *Int. J. Chem. Stud.* **2017**, *5* (2), 350–362.

(140) Townsend, L. B.; Wise, D. S. *Parasitol. Today* **1990**, *6* (4), 107–112.

(141) Phillips, M. A. *J. Chem. Soc.* **1928**, 0, 172–177.

(142) Phillips, M. A. *J. Chem. Soc.* **1930**, 0, 1409–1419.

(143) Hein, D. W.; Alheim, R. J.; Leavitt, J. J. *J. Am. Chem. Soc.* **1957**, *79* (2), 427.

(144) Preston, P. N. *Chem. Rev.* **1974**, *74* (3), 279–314.

(145) So, Y.-H.; Heeschen, J. P. K. *J. Org. Chem.* **1997**, *62* (10), 3552–3561.

(146) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Comb. Chem.* **1999**, *1* (1), 55–68.

(147) Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61* (46), 10827–10852.

(148) *J. Org. Chem.* **1961**, *26* (7), 2525–2528.

(149) Sheehan, J. C.; Hess, G. P. *J. Am. Chem. Soc.* **1955**, *77* (4), 1067–1068.

(150) Neises, B.; Steglich, W. *Angew. Chemie Int. Ed. English* **1978**, *17* (7), 522–524.

(151) Wittenberger, S. J.; McLaughlin, M. A. *Tetrahedron Lett.* **1999**, *40* (40), 7175–7178.

(152) Clayden, J., Greeves, N., Warren, S., Wothers, P. *Organic Chemistry*, 1st ed.; Oxford University Press, 2000.

(153) Li, Z.; Zhu, A.; Yang, J. *J. Heterocycl. Chem.* **2012**, *49* (Scheme 1), 1458–1461.

(154) Taber, D. F.; Tirunahari, P. K. *Tetrahedron* **2011**, *67* (38), 7195–7210.

(155) Epichlorohydrin http://www.sigmaaldrich.com/catalog/product/sial/45340?lang=en&region=ZA&cm_sp=Insite-_-prodRecCold_xorders-_-prodRecCold2-1 (accessed Oct 14, 2017).

(156) Chua, M. S.; Shi, D. F.; Wrigley, S.; Bradshaw, T. D.; Hutchinson, I.; Shaw, P. N.; Barrett, D. A.; Stanley, L. A.; Stevens, M. F. G. *J. Med. Chem.* **1999**, *42* (3), 381–392.

(157) Kimura, T.; Hosokawa-Muto, J.; Asami, K.; Murai, T.; Kuwata, K. *Eur. J. Med. Chem.* **2011**, *46* (11), 5675–5679.

(158) Hou, T.; Xu, X. *Curr. Pharm. Des.* **2004**, *10* (9), 1011–1033.

(159) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins Struct. Funct. Genet.* **2002**, *47* (4), 409–443.

(160) Neamati, N.; Barchi, J. J. *Curr. Top. Med. Chem.* **2002**, *2* (September), 211–227.

(161) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. *Med. Res. Rev.* **2006**, *26* (5), 531–568.

(162) Discovery Studio Modeling Environment, v4.0, **2013**, *Accelrys Software Inc.*, San Diego.

(163) Akkermans, R. L. C.; Spenley, N. A.; Robertson, S. H. *Molecular Simulation*. Taylor & Francis 2013, pp 1153–1164.

(164) Gilli, P.; Bertolasi, V.; Ferretti, V.; Gilli, G. *J. Am. Chem. Soc.* **1994**, *116* (3), 909–915.

(165) Janiak, C. *J. Chem. Soc. Dalt. Trans.* **2000**, No. 21, 3885–3896.

(166) Urayama, S.; Fukuyama, H.; Stockton, K.; Usui, K.; Green, D. W.; Price, C. J. *Science* **2017**, *312* (5779), 1537–1540.

(167) Makler, M. T.; Ries, J. M.; Williams, J. A.; Bancroft, J. E.; Piper, R. C.; Gibbins, B. L.; Hinrichs, D. J. *Am. J. Trop. Med. Hyg.* **1993**, *48* (6), 739–741.

(168) Combrinck, J. M.; Fong, K. Y.; Gibhard, L.; Smith, P. J.; Wright, D. W.; Egan, T. J. *Malar. J.* **2015**, *14* (1), 253, 1-14.

(169) Accelrys Software Inc.: San Diego 2013.

(170) Bevan, C. D.; Lloyd, R. S. *Anal. Chem.* **2000**, *72* (8), 1781–1787.

(171) Li, Y.; De Kock, C.; Smith, P. J.; Chibale, K.; Smith, G. S. *Organometallics* **2014**, *33* (17), 4345–4348.

(172) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2012**, *64* (SUPPL.), 4–17.

(173) Vanaerschot, M.; Lucantoni, L.; Li, T.; Combrinck, J. M.; Ruecker, A.; Kumar, T. R. S.; Rubiano, K.; Ferreira, P. E.; Siciliano, G.; Gulati, S.; Henrich, P. P.; Ng, C. L.; Murithi, J. M.; Corey, V. C.; Duffy, S.; Lieberman, O. J.; Veiga, M. I.; Sinden, R. E.; Alano, P.; Delves, M. J.; Lee Sim, K.; Winzeler, E. A.; Egan, T. J.; Hoffman, S. L.; Avery, V. M.; Fidock, D. A. *Nat. Microbiol.* **2017**, *2* (10), 1403–1414.

(174) Bellina, F.; Calandri, C.; Cauteruccio, S.; Rossi, R. *Tetrahedron* **2007**, *63* (9), 1970–1980.

(175) Ayhan-Kılcıgil, G.; Altanlar, N. *Farm.* **2003**, *58* (12), 1345–1350.

(176) Sangermano, M.; Malucelli, G.; Priola, A.; Lengvinaite, S.; Simokaitiene, J.; Grazulevicius, J. V. *Eur. Polym. J.* **2005**, *41* (3), 475–480.

(177) Sundararajan, C.; Falvey, D. E. **2005**, No. 8, 7687–7690.

(178) Harada, H.; Fujii, A.; Kato, S. *Synth. Commun.* **2003**, *33* (3), 507–514.

(179) GraphPad Prism, v4.0, **2003**, San Diego California USA, www.graphpad.com.

(180) MarvinSketch v6.2.2., **2014**, *ChemAxon Ltd*., http://www.chemaxon.com

(181) ChemDraw Ultra v12.02, **2010**, *CambridgeSoft*, www.cambridgesoft.com